

SERDP Interim Report on Downscaling Nonstationarity

Project Title:

Process-based Evaluation of Temperature and Precipitation
Projections and Downscaling Methods over the CONUS:
Charting a Path for End-Users from the CMIP6 Ensemble to
Multivariate Facility-Level Risks

SERDP Project Number:

RC19-1391

Performing Organization:

Lawrence Berkeley National Laboratory

Lead Principal Investigator:

Daniel Feldman

Date:

22 August 2022

Version Number:

2 .0

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY) 22/078/2022		2. REPORT TYPE SERDP Interim Report			3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Measuring and Predicting the Natural and Enhanced Rate and Capacity of Abiotic Reduction of Munition Constituents Determining Properly Normalized Second Order Rate Constants and Reactive Reductant Concentrations: Limited Reductant Kinetic Experiments				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Pei Chiu, Dominic M. Di Toro, Kevin P. Hickey, Jimmy Murillo Gelvez, Paula Cardenas Hernandez, Herbert E. Allen University of Delaware Richard F. Carbonaro Mutch Associates				5d. PROJECT NUMBER RC19-1391		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lawrence Berkeley National Laboratory 1 Cyclotron Rd M/S 74R316C Berkeley, CA 94720					8. PERFORMING ORGANIZATION REPORT NUMBER RC19-1391	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Strategic Environmental Research and Development Program 4800 Mark Center Drive, Suite 16F16 Alexandria, VA 22350-3605					10. SPONSOR/MONITOR'S ACRONYM(S) SERDP	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S) RC19-1391	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT There is a critical need for granular information on how surface temperature and precipitation statistical distributions will change across the Conterminous United States (CONUS) at the DoD Facility Level throughout the 21st Century. Ensuring resilience for infrastructure and operations in the face of nonstationarity, where surface conditions that occurred in the past may not encompass the range of conditions in the future, requires this information. Unfortunately, the primary tools for developing projections of future surface temperature and precipitation, Earth System Models (ESMs), are producing overwhelmingly large datasets that both disagree with each other at the regional level and do not resolve the hydroclimate processes that are most likely to impact Facilities . Therefore, the primary objective of RC19-1391 is to develop localized, accessible surface temperature and precipitation projections, along with their uncertainties, using state-of-the-science information from ESMs.						
15. SUBJECT TERMS Downscaling, nonstationarity, temperature, precipitation, CMIP6, LOCA2, WRF						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UNCLASS	18. NUMBER OF PAGES 61	19a. NAME OF RESPONSIBLE PERSON Daniel Feldman	
a. REPORT UNCLASS	b. ABSTRACT UNCLASS	c. THIS PAGE UNCLASS			19b. TELEPHONE NUMBER (Include area code) 5104952171	

Table of Contents

Front Matter	iii
List of Tables	iii
List of Figures	iv
List of Acronyms	v
Keywords	vi
Acknowledgements	vi
Abstract	1
Executive Summary	2
Introduction	2
Objectives	4
Technical Approach	4
Results and Discussion	5
Future Research Directions and Benefits to DoD	6
Full Report	7
Objective	7
Background	10
CMIP6	12
Statistical Downscaling Overview	15
Dynamical Downscaling Overview	20
Nonstationary Analysis of Downscaling Solutions	24
Parsimoniously Downscaling CMIP6	31
Observational Product Analysis	36
Model Weighting	42
Future Research Directions and Benefits to DoD	43
Literature Cited	46

Front Matter

List of Tables

1. Interim Report: Description of Shared Socioeconomic Pathways Used
2. Interim Report: List of CMIP6 Models which LOCA2 will downscale
3. Interim Report: List of WRF-downscaled CMIP6 Models
4. Interim Report: Precipitation return value error contributed by time-of-day vs. gridding
5. Interim Report: Metrics used for model weighting by 5th National Climate Assessment

List of Figures

1. Exec. Summary: Flow-chart for translating CMIP projections to local risk.
2. Interim Report: RC19-1391 Project Info-graphic.
3. Interim Report: Survey of CMIP6 temperature and precipitation data holdings on ESGF
4. Interim Report: Diagram of statistical downscaling approach
5. Interim Report: Example of statistical downscaling analog day selection
6. Interim Report: Challenges with historical observations for projecting extreme precip.
7. Interim Report: Update to extreme precipitation downscaling in LOCA2
8. Interim Report: Diagram of LOCA2 statistical-downscaling algorithm
9. Interim Report: Dynamical downscaling domain configuration
10. Interim Report: Dynamical downscaling development with observational metrics
11. Interim Report: Performance of dynamical downscaling relative to observations
12. Interim Report: Nonstationarity in statistical downscaling of temperature over California
13. Interim Report: Nonstationarity in statistical downscaling of precipitation over California
14. Interim Report: Effects of bias-correction on dynamical downscaling solutions
15. Interim Report: Difference between bias-corrected and non-bias-corrected downscaling
16. Interim Report: Upstream model error diagnostics for dynamical downscaling
17. Interim Report: Shelf-life analysis of bias-correction in downscaling
18. Interim Report: Taylor Diagram of CMIP6 ensemble for precipitation over CONUS
19. Interim Report: Taylor Diagram of CMIP6 ensemble for precipitation over US regions
20. Interim Report: Ratio of between-to-within-model variability of CMIP6 ensemble
21. Interim Report: Distribution of times-of-day for CONUS precipitation data collection
22. Interim Report: Differences in 20-year precip. return value between R19 and L15/LOCA
23. Interim Report: Correction factor for precipitation return value of L15/LOCA
24. Interim Report: Homogenized vs non-homogenized temperature trends in California

List of Acronyms

AR6:	Sixth Assessment Report
CMIP5:	Coupled Model Intercomparison Project – Phase 5
CMIP6:	Coupled Model Intercomparison Project – Phase 6
CMIP7:	Coupled Model Intercomparison Project – Phase 7
CONUS:	Conterminous United States
COSMO:	Consortium for Small-Scale Modeling
CPU:	Central Processing Unit
DECK:	Diagnostics, Evaluation, and Characterization of Klima
DOD:	Department of Defense
DTR:	Daily Temperature Range
ENSO:	El Niño Southern Oscillation
ERA5:	European Centre for Medium-Range Weather Forecasting ReAnalysis -5
ESGF:	Earth System Grid Federation
GCM:	Global Circulation Model
GDO:	Greed Data Oasis
GMT:	Global Mean Temperature
GP:	Gaussian Process
GPCP:	Global Precipitation Climatology Project
H1:	RC19-1391's first null hypothesis
H2:	RC19-1391's second null hypothesis
HEC:	High-End Computing (NASA)
HighResMIP:	High-Resolution Model Intercomparison Project
IPCC:	Intergovernmental Panel on Climate Change
L15:	Livneh et al, 2015 gridded hydrometeorological dataset
LBNL:	Lawrence Berkeley National Laboratory
LLNL:	Lawrence Livermore National Laboratory
LOCA:	Localized Constructed Analogs
LOCA2:	Localized Constructed Analogs, Version 2
MIP:	Model Intercomparison Project
MM5:	Fifth-Generation Penn State/NCAR Mesoscale Model
NA-CORDEX:	North American Coordinated Regional Downscaling Experiment
NASA:	National Aeronautics and Space Administration
NARR:	North American Regional Reanalysis
NCA4:	Fourth National Climate Assessment
NCA5:	Fifth National Climate Assessment
NOAA:	National Oceanic and Atmospheric Administration
P3:	Predicted Particle Properties
PCMDI:	Program for Climate Modeling Diagnostics and Intercomparison
PDF:	Probability Distribution Function
PRISM:	Parameter-elevation Regressions on Independent Slopes Model
R19:	Risser et al, (2019a) dataset
RCM:	Regional Climate Model
RCP:	Representative Concentration Pathway
RCP4.5:	Representative Concentration Pathway, version 4.5

RCP8.5:	Representative Concentration Pathway, version 8.5
QM:	Quantile Mapping
RMS:	Root Mean Square
SAM:	System for Atmospheric Modeling
SSP:	Shared Socioeconomic Pathway
SSP245:	Shared Socioeconomic Pathway 2 with 4.5 W/m ² radiative forcing in 2100
SSP370:	Shared Socioeconomic Pathway 3 with 7.0 W/m ² radiative forcing in 2100
SSP585:	Shared Socioeconomic Pathway 5 with 8.5 W/m ² radiative forcing in 2100
SERDP:	Strategic Environmental Research and Development Program
SNOTEL:	Snow Telemetry
SWE:	Snow Water Equivalent
TOA:	Top-of-Atmosphere
TOD:	Time-of-Day
USACE:	United States Army Corps of Engineers
USGCRP:	United States Global Change Research Program
VIACS:	Vulnerability, Impacts, Adaptation, and Climate Services
WCRP:	World Climate Research Program
WRF:	Weather Research and Forecasting
WUS:	Western United States
YSU:	Yonsei University WRF planetary boundary layer scheme

Keywords

Downscaling, nonstationarity, temperature, precipitation, CMIP6, LOCA2, WRF

Acknowledgements

The Lead Principal Investigator would like to acknowledge the team of researchers whose findings and work substantially contributed to this report including Dr. Mark Risser and Ms. Julia Longmate of the Lawrence Berkeley National Laboratory, Drs. Daniel Cayan and David Pierce of the University of California-San Diego, Drs. Stefan Rahimi-Esfarjani, Naomi Goldenson, and Alex Hall and Mr. Will Krantz of the University of California-Los Angeles, Drs. Travis O'Brien and Alex Charn of Indiana University, and Dr. Jeffrey Arnold previously of the United States Army Corps of Engineers (USACE) and now at MITRE. This team would also like to recognize the detailed discussions with Dr. Michael Wehner of the Lawrence Berkeley National Laboratory who provided valuable feedback on the Coupled Model Intercomparison Project (CMIP) landscape and extrema analysis, Dr. Ken Kunkel of North Carolina State University, Dr. Dave Easterling of the National Oceanic and Atmospheric Administration, Dr. Fred Lipschultz and Dr. Samantha Basile for suggestions and feedback on where this project's deliverables would contribute to the Fifth National Climate Assessment, and Drs. Kyo Lee of the NASA Jet Propulsion Laboratory, Keith Dixon of NOAA Geophysical Fluid Dynamics Laboratory, and Elias Massoud of the NASA Jet Propulsion Laboratory and the University of California—Berkeley for guidance and discussion on where this project's deliverables could contribute to CMIP6 model weighting efforts. Finally, the investigating team would like to acknowledge the SERDP program for providing the funding to perform this research through Project RC19-1391: “Process-based Evaluation of Temperature and Precipitation Projections and Downscaling Methods over the CONUS: Charting a Path for End-Users from the CMIP6 Ensemble to Multivariate Facility-Level Risks.”

Abstract

Introduction and Objectives:

There is a critical need for granular information on how surface temperature and precipitation statistical distributions will change across the Conterminous United States (CONUS) at the DoD Facility Level throughout the 21st Century. Ensuring resilience for infrastructure and operations in the face of nonstationarity, where surface conditions that occurred in the past may not encompass the range of conditions in the future, requires this information. Unfortunately, the primary tools for developing projections of future surface temperature and precipitation, Earth System Models (ESMs), are producing overwhelmingly large datasets that both disagree with each other at the regional level and do not resolve the hydroclimate processes that are most likely to impact Facilities. Therefore, the primary objective of RC19-1391 is to develop localized, accessible surface temperature and precipitation projections, along with their uncertainties, using state-of-the-science information from ESMs.

Technical Approach:

This project focuses first on analyzing the ~40 Petabytes of ESM simulations from the Coupled Model Intercomparison Project – Phase 6 (CMIP6) by evaluating Earth System Model skill relative to historical observations. It then develops statistical and dynamical downscaling solutions to produce surface temperature and precipitation at 6 km resolution using Localized Constructed Analogs Version 2 (LOCA2) and the Weather Research and Forecasting (WRF) model, respectively. Finally, the project determines where nonstationary effects influence surface temperature and precipitation and produces a dataset with uncertainty estimates for end-users.

Results:

To date, the RC19-1391 research has produced terabytes of statistical and dynamical downscaling data. From these, it has produced 5 publications with interim scientific findings that show a defensible path forward to downscaling CMIP6 for end-users. Findings include:

1. Uncertainties in historical observations greatly complicate the assessment of model skill.
2. CMIP6 downscaling should focus on multiple models, not multiple ensemble members.
3. Earth system models should be bias-corrected before any downscaling is performed.
4. The influence of bias-correction on future downscaled projections appears to be linear.

For the remainder of the RC19-1391 project, it will focus on analyzing LOCA2 and WRF downscaled datasets to ensure that 21st Century projections of temperature and precipitation do not underestimate risk to DoD Facilities.

Benefits:

This research is developing robust, scientifically-based datasets for climate vulnerability assessments with state-of-the-science downscaling that directly address the findings of ESTCP Project RC-201591. These can be widely utilized by the scientific and DoD user communities.

Executive Summary

Introduction

According to [Milly et al. \(2008\)](#) and as quoted in the Climate Change Installation Adaptation and Resilience 2017 Report, “stationarity is dead,” and “In the new reality of non-stationarity, projections about climate must assume a constantly evolving basis.” But how does this translate practically to making critical infrastructure and planning decisions?

Large infrastructure and operations planning decisions are strongly influenced by average surface temperature and precipitation, and variability in these quantities including extreme events. Such decisions may require many years of lead-time. Potential changes in the distributions of surface temperature and precipitation confound infrastructure and planning resiliency: for example, a straightforward analysis of distributions of those quantities based historical observations, which was a standard practice ([NRC, 1982](#)), may not account for the risks that the infrastructure will be facing over its lifetime. Similarly, changes to operations plans may be required as surface temperature and precipitation exhibit features without historical analogues.

Earth System Models (ESMs) represent the primary scientific tool for developing projections of change in the Earth system in the 21st Century that are expected to occur due to changing atmospheric composition. These global climate models contain a vast amount of the scientific community’s understanding of the Earth system and its responses to forcing from increasing greenhouse gases, aerosols, and land-use change. They are invaluable to advancing the scientific understanding of the Earth system and, increasingly, to plan for future climatic conditions outside of the scientific arena. They can be used to assess changes in the large-scale distribution of temperature and precipitation as a function of future human activities ([Hawkins and Sutton, 2009](#)). They do so by capturing temperature and precipitation change at regional scales through thermodynamical and dynamical responses to the radiative forcing from greenhouse gases, which are conventionally divided into discrete greenhouse gas emissions scenarios ([van Vuuren et al, 2011](#)). These models are complex numerical representations of conservation equations of mass, energy and momentum, and consequently exhibit a wide range of Earth system phenomena that are also observed. For example, the response of the Earth’s hydrological cycle as it adjusts to warming ([Jeevanjee and Romps, 2018](#)), the expansion of the troposphere ([Lu et al, 2007](#); [Hu et al, 2013](#)), and shifts in the storm tracks ([Yin, 2005](#); [Bengtsson et al, 2006](#); [Ulbrich et al, 2008](#)), have been both observed and modeled, while also having a strong theoretical basis.

Still, the ESMs, because they only approximate the conservation equations, disagree with each other in their projections of changes in temperature ([Charney et al, 1979](#); [Knutti and Hegerl, 2008](#); [Forster et al, 2016](#); [Knutti et al, 2017](#)) and precipitation in the future. The disagreements arise due to a combination of parameterization uncertainties, model component interaction uncertainties, and, in some cases, model errors.

While there has been considerable scientific focus on narrowing the range of global-scale temperature and precipitation responses in ESMs to increased atmospheric carbon dioxide concentrations, infrastructure and operations resiliency require details of what local impacts are likely to occur as the result from global-scale changes, so ESM uncertainty for particular end-users must be managed.

Another unfortunate complication to using ESMs for planning purposes is that they are coarse-resolution and therefore blunt instruments for projecting local climate change. The spatial resolution of ESMs is typically ~100 km. While higher resolution ESM simulations are quite possible (Haarsma et al, 2016), they are too computationally expensive for the modeling centers because those centers are using ESMs to perform a wide range of experiments of centennial and longer timescales. With such coarse ESM resolution, the processes and details that are necessary to resolve the effects that influence surface temperature and precipitation at the DoD Facility scale are simply not included in ESMs. Additional information on local processes is required.

With ESMs producing a range of large-scale projections while also not resolving the details necessary for climate change impacts analysis at the DoD facility scale, there is no obvious path to using ESMs directly for resilient infrastructure and/or operations planning to explicitly account for nonstationarity effects. For infrastructure especially, there is a fundamental gap between the need to make infrastructure decisions conservatively based on observations of actual conditions that infrastructure has faced and the recognition that those observations may not be sufficiently representative of conditions that infrastructure will face in the near future.

This tension has led to growth of resiliency research and planning to fill this gap. It also led the Strategic Environmental Research and Development Program (SERDP) to issue a Statement of Need entitled “Quantifying the Effects of Environmental Nonstationarity at DoD Relevant Scales” (RCSO-18-L2). This tension also led to the convening of a scientific workshop supported by the Environmental Security Technology Certification Program (ESTCP) entitled “Nonstationary Weather Patterns and Extreme Events Informing Design and Planning for Long-Lived Infrastructure” (ESTCP Project RC-201591) which principally found that the current level of scientific understanding has room for improvement and that practitioners require additional information on (1) a wide range of hazards and thresholds; (2) the most likely conditions and maximum credible extremes for a number of periods and mean recurrence intervals; (3) historical and projected conditions analysis at high temporal and spatial resolution for specific sites; and (4) considerations for requirements sensitive to practitioner discipline and location.

The RC19-1391 project “Process-based Evaluation of Temperature and Precipitation Projections and Downscaling Methods over the CONUS: Charting a Path for End-Users from the CMIP6 Ensemble to Multivariate Facility-Level Risks” responds to these findings by focusing on developing localized climate projections from the Coupled Model Intercomparison Project – Phase 6 (CMIP6) (Eyring et al, 2016). It builds off research supported by RCSO-18-L2 and the ESTCP Project RC-201591 workshop recommendations to develop robust estimates of changes, and associated confidence intervals, in surface temperature and precipitation distributions at the DoD facility level across the Conterminous United States (CONUS).

This interim report delivers updates on the findings of the RC19-1391 project to date. Through these updates, this report provides context for the project's research activities and its contributions to advancing the scientific understanding of downscaling over the CONUS. Furthermore, it highlights both the opportunities and the challenges of developing robust downscaling of surface temperature and precipitation in the face of myriad interacting processes which produce a range of nonstationary effects. The report concludes with a high-level finding: the downscaling research in this project should, as it proceeds, seek to identify and address the sources nonstationarity in the downscaling of CMIP6 over the CONUS that can or will lead to an underestimation of risk to infrastructure from changing temperature and precipitation distributions.

Objectives

The primary objective of this project is to develop localized projections of changes in surface temperature and precipitation across the CONUS in the 21st Century that summarize and distill state-of-the-science ESM projections and downscaling solutions. This will add substantially to the foundation of climate information that will help build a comprehensive determination of the risk to DoD infrastructure and operations from expected changes in temperature and precipitation.

Technical Approach

This research focuses first on developing and producing downscaled surface temperature and precipitation solutions, and next on analyzing both ESM output and downscaling solutions to curate these large datasets and provide distilled information for end-users. The research focuses on three scenarios: a business-as-usual scenario (RCP8.5/SSP585), a scenario with slightly-mitigated policies (RCP7.0/SSP370) and a scenario with mitigated climate change effects (RCP4.5/SSP245).

Because the purpose of the generation of downscaled projections of surface temperature and precipitation is to develop state-of-the-science datasets that capture any expected nonstationary climatic effects, the project is organized around the testing of two null hypotheses. The first of these (H1) is that, for the purposes of CONUS downscaling, each model is an independent, unbiased sample of the trajectory of Earth's climate in the 21st Century. In order to test this hypothesis, this research is critically evaluating ensembles of CMIP6 model simulations over the CONUS against historical observations and observational products to determine both how skillful each model is and how independent it is from other models.

The second null hypothesis (H2) is that downscaling methods, trained and tested on historical observations, are appropriate and unbiased for the production of localized projections of temperature and precipitation fields in the 21st Century over the CONUS. This hypothesis focuses on ensuring that downscaling solutions translate the large-scale ESM forcing to the local-scale in ways that accurately assess risk.

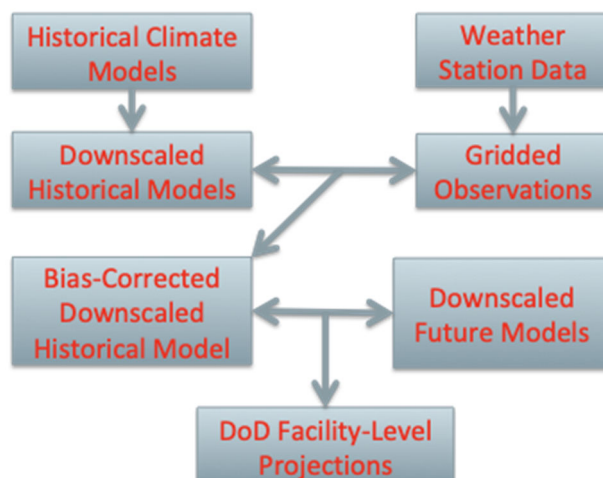


Figure 1: Flow-chart indicating how CMIP model projections are translated into model facility-level risk. Adapted from Feldman et al, (2021).

Figure 1 shows the process chain for developing downscaling solutions: parent ESM simulations are compared against gridded observational data products over the historical period to assess model skill and independence and also to produce estimates of bias-correction needs. Model skill over the historical period across a range of metrics is used to develop downscaled projections of future surface temperature and precipitation that can be further processed and analyzed to assess DoD facility level risks in the 21st Century.

This process chain, while seemingly straightforward, is exceedingly difficult to navigate because of the myriad assumptions inherent in each step and the mismatch between the capabilities of ESMs and end-user needs for realistic, accurate projections.

Results and Discussion

The project has undertaken two substantial downscaling efforts: it has updated the Localized Constructed Analogs (LOCA) statistical downscaling methodology to produce LOCA2. That capability has downscaled 27 CMIP6 models and has provided those results to the 5th National Climate Assessment (NCA5). Additionally, centennial-length dynamical downscaling has been developed with WRF over the Western United States. These efforts are both central to the project's objective and are prerequisites for testing H2.

The project has also navigated the 20-40 Petabyte CMIP6 archive and developed methods to sample that archive parsimoniously for downscaling purposes. Additionally, the project has developed advanced methods to test model performance against observational datasets. These efforts are central to testing H1.

Finally, the project has started activities to dovetail with the testing of these two null hypotheses: based on a careful and comprehensive analysis of ESMs and downscaling solutions using process-based metrics, and in coordination with model weighting activities of the Fifth National Climate Assessment (NCA5), the project is focused on a final task of producing a distilled set of

surface temperature and precipitation projections over the CONUS in the 21st Century and, most importantly, producing defensible estimates of their uncertainties.

Several publications have been produced by this project including the following:

1. Charn et al, (2022) found that the accuracy of the observational products themselves must be first established before assessing model skill (H1)
2. Pierce et al, (2021) and Risser et al (2021) showed the updates needed for LOCA2 (H2).
3. Longmate et al (In Revision) showed how to sample CMIP6 parsimoniously for downscaling purposes (H1).
4. Rahimi et al (2022) showed how the configuration for an unprecedented set of dynamical downscaling activities that have been completed for CMIP6 (H2).

Future Research Directions and Benefits to DoD

This interim report discusses a set of research activities that are in-progress, but for which a substantial amount of legwork has been performed. The project will be focused on completing its tasks and objectives during its stated period of performance (12/31/2019 to 12/31/2023).

The latter half of this project will focus on process-based analysis of LOCA2 and WRF downscaling solutions and will focus that analysis on those locations where different statistical and/or dynamical solutions diverge, since that divergence points to the sensitivity of the solution to different downscaling assumptions and the potential role of nonstationary effects. More focus on the shelf-life of downscaling solutions and assumptions therein will be undertaken. The project will focus the analyses of divergence on ensuring that risk is not underestimated due to stationarity assumptions.

Finally, the project will use this set of investigations to meet its primary objective and deliver a distilled, easy-to-use, and defensible set of downscaled temperature and precipitation fields over the CONUS and their uncertainties to end-users, wherein uncertainties from nonstationary effects are explicitly estimated.

This research has developed robust, scientifically-based datasets for climate vulnerability assessments with state-of-the-science downscaling that directly address the findings of ESTCP Project RC-201591. These results have been delivered and analyzed by Chapter Authors in the 5th National Climate Assessment and can be widely utilized by the scientific and DoD user communities to enable planning for surface temperature and precipitation nonstationarity impacts on infrastructure and operations.

Full Report

Objective

Because DoD infrastructure and operations are directly impacted by surface temperature and precipitation conditions, planning across the Department that mitigates and minimizes impacts of those conditions on infrastructure and operations is needed. That planning has been and needs to continue to be fully cognizant of the range of surface temperature and precipitation conditions at the many locations where its infrastructure and operations activities occur. Therefore, there are ongoing needs for the best available information on those conditions now and throughout the 21st Century (e.g., Office of the Under Secretary of Defense for Acquisition and Sustainment, 2019).

Unfortunately, assuming that the conditions of the past, where there is ample information to inform decision-making, will simply persist into the future is untenable (Milly et al, 2008). The changes in atmospheric composition, particularly the rise of atmospheric carbon dioxide, and the response of the Earth system and its energy budget that have occurred and are projected to occur in the 21st Century have already produced and are likely to continue to produce secular changes in myriad aspects of the Earth's temperature and precipitation patterns. They fundamentally challenge the assumption of stationarity in the statistical distributions of temperature and precipitation at the time and space scales relevant for DoD infrastructure and operations planning.

Therefore, this research has the primary objective of **developing comprehensive, localized, and distilled datasets that provide end-users with state-of-the-science estimates of the coming climate change effects on surface temperature and precipitation in the 21st Century over the Conterminous United States (CONUS).**

Figure 2 shows an infographic of the steps that this project is taking to achieve this objective, and these directly map onto tasks that the project is fulfilling (described below). The Figure starts with a range of different emissions scenarios that comprise a set of different pathways that humanity may undertake in the 21st Century. These pathways range from highly-mitigated emissions (Shared Socioeconomic Pathway 1) to business-as-usual emissions (Shared Socioeconomic Pathway 5), which are constructed by considering population, technology, and economic factors (O'Neill et al, 2014; 2016). These emissions scenarios are forcings for climate models because it has been consistently demonstrated that surface temperature and precipitation patterns of the 21st Century will depend, at multi-decadal time-scales, on anthropogenic emissions of greenhouse gases, aerosols, and land-use changes (Hawkins and Sutton, 2009).

These emissions pathways serve as inputs to the Earth System Models (ESMs) that contributed results to the Coupled Model Intercomparison Project – Phase 6 (CMIP6). Figure 2 depicts the CMIP6 models as a set of discretized simulations of Earth. Because the ESMs are too coarse for localized projections (~100 km) resolution, the CMIP6 ESMs are subjected to two complementary downscaling approaches that produce daily surface temperature and precipitation projections at 6 km resolution or finer.

21st Century Temperature and Precipitation Downscaling Projections (RC19-1391)

*How will climate change impact local temperature and precipitation patterns at the scale of military installations?
Which climate data products should be used when making long-term decisions?*

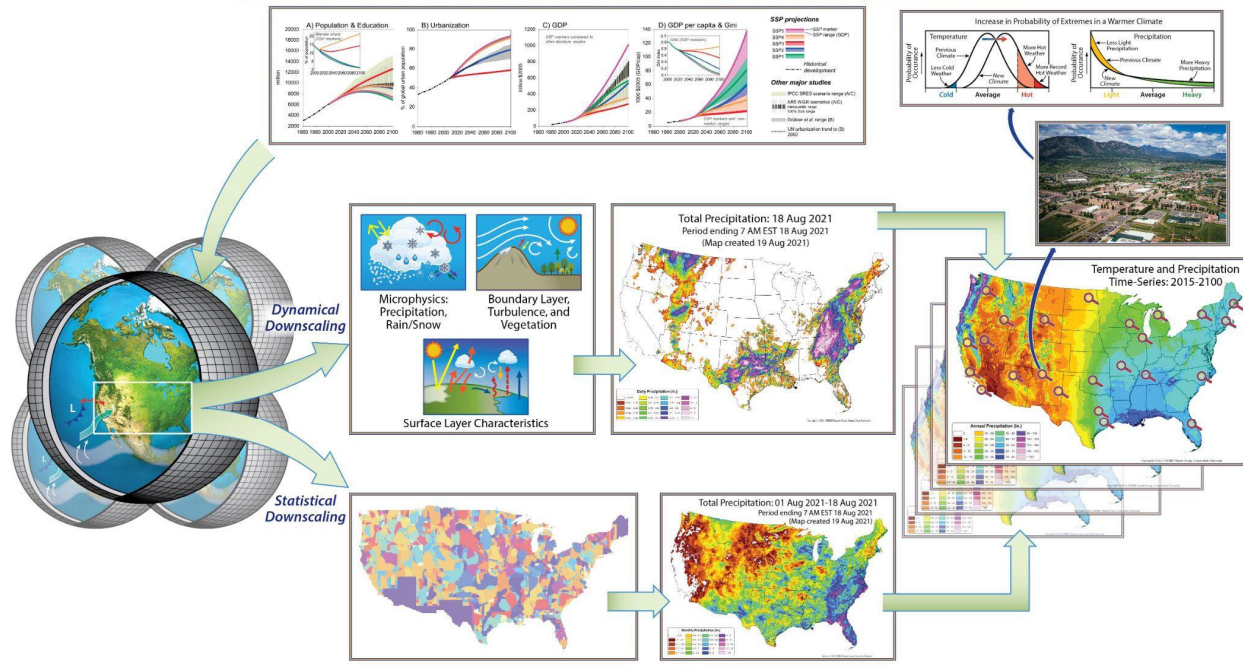


Figure 2: Flow diagram of the dynamical and statistical downscaling activities that RC19-1391 develops to analyze and assess changes in surface temperature and precipitation distributions at the DoD facility scale in the 21st Century depending on emissions scenarios.

Figure 2 shows that dynamical downscaling methods is one of the approaches used in this project. This type of downscaling consists of using numerical simulations of mass, momentum, and energy at spatial and temporal resolutions that are much finer than those of ESMs (typically as fine as 250 m, but can be as fine as tens of meters for Large Eddy Simulations). Dynamical downscaling simulations consist of atmospheric process models (such as the Weather Research and Forecasting (WRF) model, the Consortium for Small-Scale Modeling (COSMO), or the System for Atmospheric Modeling (SAM)) which represent the dominant atmospheric and land-atmosphere-interaction physical processes that impact surface temperature and precipitation. As shown in Figure 2, they simulate the atmospheric circulation and capture its evolution in a spatial domain of interest when forced by an ESM. The atmospheric process models produce highly-detailed maps of atmospheric variables including surface temperature and precipitation at sub-daily resolution over the 21st Century. Because of the explicit representation of the governing physical processes that impact surface temperature and precipitation in dynamical downscaling, such approaches can model those variables under conditions that are without any historical analogue, thereby capturing potential nonstationary climate effects

However, dynamical downscaling is computationally expensive and still does not fully resolve the processes that impact surface temperature and precipitation. Since space and timescales of millimeters and seconds, respectively, that are needed to fully resolve these processes, atmospheric process models still contain parameterizations of key processes that impact the variables of interest. Therefore, statistical methods, which are built explicitly from observational

datasets, represent a key alternative. Figure 2 shows how analog-based statistical downscaling identifies a large set of regional-scale weather patterns in the historical record where the historical atmospheric circulation closely resembles the atmospheric circulation projected by the ESM. The surface temperature and precipitation at local scales are the result of many processes that are governed by that large-scale circulation, which are not fully resolved in dynamical downscaling but are captured in statistical relationships in the historical record. Statistical downscaling, therefore, presents an important downscaling solution that contrasts with dynamical downscaling, and even though such methods assume that the relationships between the large-scale circulation and local impacts are stationarity.

Finally, this project collects a wide range of statistical and dynamical downscaling solutions and distills the results based on the skill of the ESMs and the downscaling approaches in order to create a compact dataset, with confidence intervals, of surface temperature and precipitation estimates at high temporal (daily) and high spatial (6 km) resolution for several different emissions scenarios over the 21st Century across CONUS. Figure 2 also depicts how statistical and dynamical downscaling results are analyzed to evaluate the changes in temperature and precipitation distributions at the facility level.

Underlying all of these efforts are this project's two null hypotheses. The first null hypothesis (H1) is that each CMIP6 simulations represent an independent, unbiased and equally plausible representation of the Earth's climate over the CONUS in the 21st Century. Since there is a large number of different global-scale ESM projections of 21st Century climate, this hypothesis concerns how researchers and end-users should manage this archive with local impacts in mind. Models that are independent of each other will produce distinct information about the importance of different nonstationary climate effects for impacts analysis, while models that are unbiased will produce estimates that enable the development of meaningful statistics on future risk. Alternatively, models that are not independent of each other will produce redundant information and can potentially mislead impacts analysis by creating the appearance of an elevated likelihood of one or more nonstationary climate effects, even though these are actually the result of multiple, dependent model simulations. Additionally, models that are biased also have the potential to mislead impacts analysis with erroneous predictions. The results of the testing of this hypothesis have direct bearing on how the CMIP6 archive and other multi-model ensemble archives (e.g., CMIP7) are downscaled. A confirmation of this hypothesis indicates that exhaustive efforts are needed to downscale CMIP6, while a rejection points to the need for a closer and more comprehensive downscaling effort focused on the CMIP6 data to capture the full range of temperature and precipitation projections of that model.

The second null hypothesis (H2) is that stationarity assumptions in statistical downscaling produce negligible errors in those downscaling solutions. This hypothesis, if confirmed, indicates that statistical methods are well-positioned to serve as the key source for localized projections of surface temperature and precipitation, and that no further adjustments are needed to those methods. However, if this hypothesis is falsified, more research and development are needed to emulate the processes that contribute to significant nonstationary effects in surface temperature and precipitation projections. That is, wholesale updates to statistical downscaling methods or even the massive adoption of dynamical downscaling methods, in spite of their computational costs, would be needed if statistical approaches need to capture these effects.

Consequently, there is a number of tasks that this project has undertaken or is in the process of undertaking to achieve this objective. They include:

1. Testing stationarity assumptions in statistical downscaling by evaluating how solutions change when observational datasets are withheld from the downscaling training.
2. Evaluating when and where statistical and dynamical solutions diverge to identify where nonstationarity could be most significant.
3. Developing and validating an updated statistical downscaling method called Localized Constructed Analogs (LOCA) and applying it to as many CMIP6 models as possible.
4. Developing process-based metrics to establish how the many CMIP6 statistical and dynamical downscaling solutions being developed should be weighted.
5. Producing a single dataset of daily surface temperature and precipitation projections across the CONUS at 6 km resolution, along with confidence intervals, for the 21st Century for 3 different emissions scenarios, to be readily utilized by end-users.

By focusing on CONUS precipitation, it therefore directly responsive to the research objective of “Discernment of relationships (e.g., mean temperature and heat wave frequency or severity) as revealed in both historical observations and climate models with an emphasis on describing model heterogeneity” as stated in SERDP SON Number: RCSON-18-L2. This research also responds to the specific research objective stated in SERDP SON Number: RCSON-18-L2 of “determining if and to what extent, mean and extreme temperature shift relationships may inform DoD on extreme risks under future scenarios or an assessment of the extent that downscaling results is needed to produce meaningful assessment”

This Interim Report is, therefore, organized as follows: background material is presented which motivates the need for downscaling climate models followed by a description of the recent international effort to develop climate model projections of the 21st Century. Next, the Report provides an overview of statistical downscaling and discusses the recent achievements by this project to advance statistical downscaling. The Report subsequently provides an overview of dynamical downscaling and also discusses the recent achievements by this project on that topic. The Report then includes findings on comparing and contrasting statistical and dynamical downscaling solutions to identify where non-stationary effects could be significant, and presents material on how to sample a large ensemble of climate model projections parsimoniously. Because downscaling solutions (both dynamical and especially statistical) are highly sensitive to the observational data products used to develop them, the project is ensuring that observational products are being used appropriately for downscaling. Finally, the project looks closely at how models are to be weighted in CMIP6. These weights will be used to produce a distilled downscaling dataset for end-users. The Report concludes with a summary of findings and research directions for the remainder of the project.

Background

Large infrastructure and operations planning decisions are made based on average surface temperature and precipitation, their variability, extreme events and their sequencing (Raymond et al, 2020). Such decisions may require many years of lead-time. Potential changes in the

distributions of temperature and precipitation confound infrastructure resiliency planning especially: a straightforward analysis of historically-observed distributions of those quantities based on data collected at the scales of interest to the infrastructure in question, which was a standard practice (NRC, 1982), may not account for the risks that the infrastructure will be facing over its lifetime (Milly et al, 2008).

The primary scientific tool for assessing changes in the Earth system due to changing atmospheric composition, such as with increasing greenhouse gases and natural variability, are Earth System Models (ESMs). These global climate models contain a vast amount of the scientific community's understanding of the Earth system and its responses to radiative forcing from increasing greenhouse gases, aerosols, and land-use change. They are invaluable to advancing the scientific understanding of the Earth system and, increasingly, to plan for future climatic conditions outside of the scientific arena. One of their increasingly common use-cases is to assess changes in the Earth System as a function of future human activities which lead to alterations in surface temperature and precipitation, which are conventionally divided into discrete greenhouse gas emissions scenarios (van Vuuren et al, 2011).

ESMs robustly capture the large-scale pattern changes of temperature and precipitation because they capture global-scale Earth system responses to the Earth's changing radiative balance. Consequently, these models exhibit a relatively narrow range of global-scale temperature responses to anthropogenic forcings (Charney et al, 1979; Knutti and Hegerl, 2008; Forster et al, 2016; Knutti et al, 2017). The response of the Earth's hydrological cycle as it adjusts to warming has also been studied, and gross features, such as the global hydrological cycle response to warming (Jeevanjee and Romps, 2018), the expansion of the troposphere (Lu et al, 2007; Hu et al, 2013), and shifts in the storm tracks (Yin, 2005; Bengtsson et al, 2006; Ulbrich et al, 2008), have been observed and have a strong theoretical basis.

Unfortunately, at the regional level ESMs models exhibit mean-state biases. These arise because they are not constructed to prioritize unbiased, high-resolution climate projections. Rather, they focus first on achieving realistic large-scale climate projections by balancing a number of competing regional-to-global-scale factors including, but not limited to, top-of-atmosphere energy balance, large-scale circulation, cloud formation and dissipation, ocean dynamics, and earth system response (Mauritsen et al, 2012). Furthermore, model development often focuses on achieving unbiased assessments of mean temperature and precipitation at regional-to-global spatial-scales and multi-decadal time-scales, with competence at synoptic spatial scales and interannual or shorter temporal scales arising as an outgrowth of a robust ESM structure. Because sub-regional variability in precipitation and precipitation change across models is quite broad (Pendergrass and Hartmann, 2014), it is not known *a priori* that a given model will produce skillful temperature and precipitation projections as they relate to DoD facilities.

ESMs also are blunt instruments for projecting local climate change impacts. The typical ESM resolution is ~100 km to account for the requirements that ESMs perform multiple centennial-length, whole-Earth simulations every few years to provide assessments (e.g., IPCC, 2021) for the scientific community to ensure that plans are developed with the potential for climate change impacts considered. The processes and details that are necessary to resolve effects that influence

temperature and precipitation at the DoD Facility scale are simply not included in ESMs. Additional information on local processes is required.

Nominally, model intercomparison projects serve to characterize the spread of model results associated with a given emissions scenario, but it has been recently recognized that navigating through the multi-model ensemble to understand the information contained in the ensemble's projections must consider both model errors with respect to observations and the interdependence of models (Sanderson et al, 2015; Knutti et al, 2017; Sanderson et al, 2017; Eyring et al, 2019). This navigation depends on the specific question being asked. Model democracy, where each simulation is treated as independent and unbiased, has been found to be suboptimal (Sanderson et al, 2015). Rather, it is essential that models being considered exhibit a level of skill with respect to regional metrics (Rupp et al, 2013). Taking California as an example, only models that reproduce the correlation and variance of mean seasonal spatial patterns, amplitude of seasonal cycle, diurnal temperature range, annual- to decadal-scale variance, long-term persistence, and Western United States regional precipitation teleconnections to El Niño Southern Oscillation (ENSO) are appropriate for hydroclimate studies (Pierce et al, 2018; Thorne et al, 2018). Only 10 of the 33 CMIP5 models exhibited skill with respect to the latter metric (ibid).

To produce regional-scale estimates of change in average temperature and precipitation from the CMIP6 ensemble, the 4th National Climate Assessment (NCA4) (USGCRP, 2017) adopted a model-weighting approach. The results broadly indicate high confidence in predicting more warming with the Representative Concentration Pathway 8.5 (RCP8.5) scenario than the RCP4.5 scenario across the Conterminous United States (CONUS), and, with less confidence, more variability in regional precipitation for both of these scenarios.

CMIP6

The Coupled Model Intercomparison Project - Phase 6 (CMIP6) has undertaken a wide breadth of experiments using Earth System Models to advance the scientific understanding of the Earth's atmosphere, its surface, and its subsurface using ESMs (Eyring et al, 2016). One of the experiments contained within CMIP6 involves different emissions scenarios as a way to sample, through scenario generation, the range of climate change of the 21st Century which depends strongly on the anthropogenic emissions (Hawkins and Sutton, 2009). This Scenario Model Intercomparison Project (ScenarioMIP) uses the Shared Socioeconomic Pathways (O'Neill et al, 2014; O'Neill et al, 2016) to capture 5 emissions pathways ranging from unmitigated to highly mitigated emissions relative to present-day conditions. These have close heritage with the Representative Concentration Pathways (van Vuuren et al, 2011) of CMIP5 (Taylor et al, 2012).

The CMIP6 archive holdings include ScenarioMIP, and as of this writing, contains terabytes of climate model simulations for that particular experiment and 20 and 40 Petabytes of simulations across all experiments. This is far too voluminous for end-users to navigate for planning purposes, and therefore there is an evident need to curate these results to determine the ranges of conditions that are likely to be faced locally. Other experiments including the North-American Coordinated Regional Downscaling Experiment (NA-CORDEX) (Giorgi et al, 2009; Jones et al, 2011; Mearns et al, 2012), the High Resolution Model Intercomparison Project (HighResMIP)

(Haarsma et al, 2016), and the Vulnerability, Impacts, Adaptation, and Climate Services (VIACS) (Ruane et al, 2016) will yield simulations at the spatial resolution of interest to DoD facilities. But again, it is expected that there will be many participating modeling centers in each of these intercomparison projects, and there is a need to determine how to interpret the range of results that arise thereof specifically in terms of DoD facility-level risk.

The flow diagram in Figure 2 depicts the approach that this project takes to curating CMIP6 projection information for DoD end-users by developing estimates of the changes in temperature and precipitation distributions at the DoD Facility Level in the 21st Century. The driving force for changes in global-scale temperature and precipitation arises from changes in the gaseous and condensate (clouds and aerosols) composition of the Earth’s atmosphere and the state of the land- and ocean-surface. These emissions scenarios drive the coarse (~100 km) resolution models that reported results to ScenarioMIP to produce projections of the Earth’s climate.

Increasing greenhouse gasses, changes in aerosols, and land-use/land-cover have increased and are projected to continue to increase the net radiative energy in the Earth’s atmosphere. In general, the addition of energy to Earth’s surface-atmosphere system will lead to increased temperatures, but many other changes in the Earth’s environment have occurred and are projected to occur from the increased radiative forcing from the changing atmospheric and surface constitution (Szopa et al, 2021).

For future projections, the dominant source of uncertainty in the 21st Century, and in particular from the 2050s onwards, is human activity (Hawkins and Sutton, 2009; Lehner et al, 2020). To isolate the social-economic-political dynamics from Earth system projections, the framework for this analysis starts with a small number of scenarios. For CMIP6, these are divided into Shared Socioeconomic Pathways (SSPs) (O’Neill et al, 2014) including ranges in population, land-use/land-cover change, and anthropogenic emissions activities associated with each SSP. The SSPs range from business-as-usual activities, whereby there is little to no climate policy mitigation (SSP5) to highly-mitigated emissions and anthropogenic activity (SSP1). Each model from CMIP6 that is participating in the Scenario Model Intercomparison Project (ScenarioMIP) (O’Neill et al, 2016) has produced both historical simulations and simulations of several SSPs extending from 1850-2100. For this project we will be evaluating three principal emissions scenarios, which are described in Table 1.

We are focusing on historical and SSP simulations reported to the CMIP6 project. Unfortunately, participation in ScenarioMIP was optional for modeling centers seeking to participate in CMIP6. While ScenarioMIP was officially endorsed by the CMIP6 panel, contributing modeling centers were only required to produce simulations associated with the Diagnostic, Evaluation and Characterization of Klima (DECK) experiments and CMIP6 historical simulations. Additionally, there were no established deadlines for the submission of simulations to the CMIP6 Earth System Federation Grid (Cinquini et al, 2014), though there were submission deadlines for model simulation participation in publications and inclusion in assessment reports such as the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (AR6).

<i>Scenario</i>	<i>Description</i>	<i>Rationale</i>	<i>Reference: DOI</i>
-----------------	--------------------	------------------	-----------------------

SSP245	SSP-based RCP scenario with medium radiative forcing by the end of the century. Following approximately RCP4.5 global forcing pathway with SSP2 socioeconomic conditions. Radiative forcing reaches a level of 4.5 W/m ² in 2100. Concentration-driven.	The scenario represents the medium part of the range of plausible future pathways.	10.5194/gmd-10-2247-2017
SSP370	Gap-filling baseline scenario with a medium to high radiative forcing by the end of century. Following approximately RCP7.0 global forcing pathway with SSP3 socioeconomic conditions. Radiative forcing reaches a level of 7.0 W/m ² in 2100. Concentration-driven.	The scenario represents the medium to high end of plausible future pathways. SSP3-7.0 fills a gap in the CMIP5 forcing pathways that is particularly important because it represents a forcing level common to several (unmitigated) SSP baselines.	10.5194/gmd-9-3461-2016
SSP585	SSP-based RCP scenario with high radiative forcing by the end of century. Following approximately RCP8.5 global forcing pathway with SSP5 socioeconomic conditions. Concentration-driven.	The scenario represents the high end of plausible future pathways. SSP5 is the only SSP with emissions high enough to produce the 8.5 W/m ² level of forcing in 2100.	10.5194/gmd-10-2247-2017

Table 1: The three emissions scenarios representing medium, medium-high, and high-emissions for the 21st Century that are used in this project for investigating nonstationarity in CMIP6 downscaling. From WCRP description in ESGF.

For ScenarioMIP simulations, not all submissions were amenable to the development of downscaling solutions for temperature and precipitation fields: if they did not provide the necessary information on the fields to be downscaled, then this project did not and will not be able to consider them for analysis. Daily precipitation, maximum temperature, minimum temperature, humidity and wind fields are all necessary variables for this downscaling project. Figure 3, from Longmate et al, (In Revision), highlights the CMIP6 data holdings and total number of initial-condition ensemble members (Kay et al, 2015; Maher et al, 2021) for each simulation as of October 11, 2021.

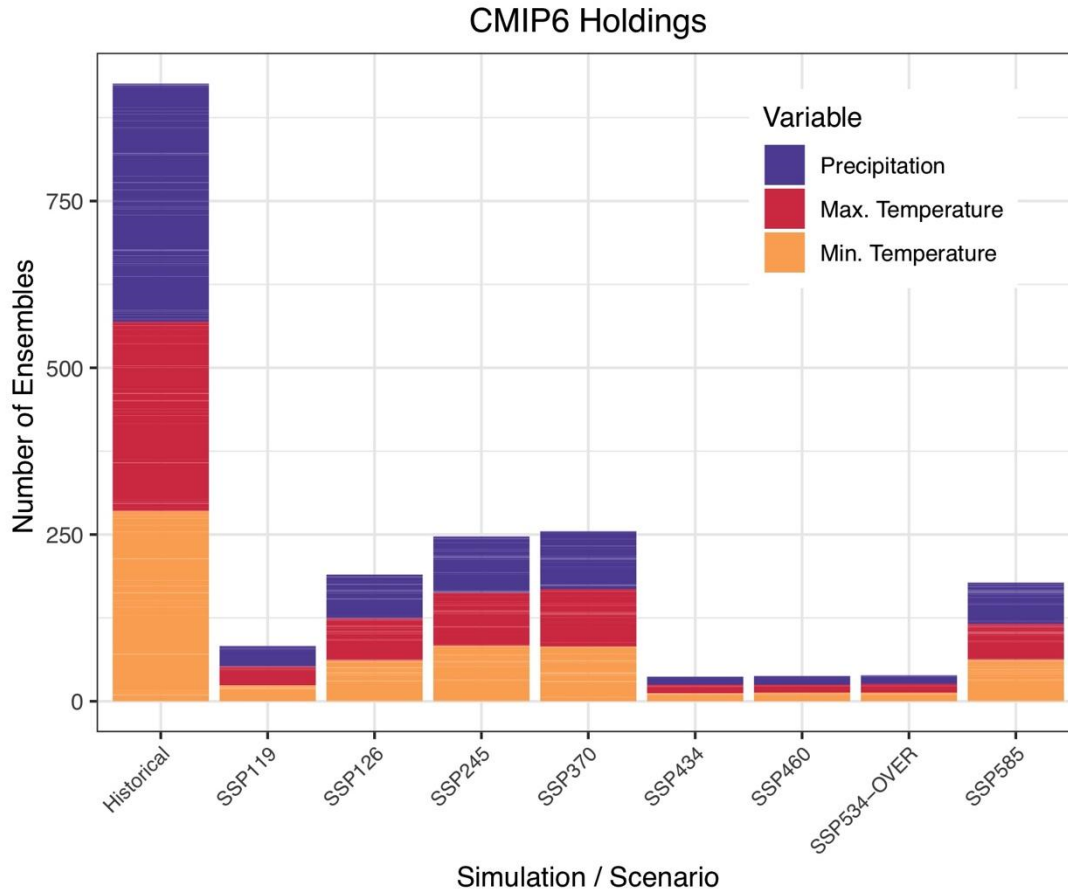


Figure 3: A survey of the CMIP6 data holdings in the ESGF as of October 11, 2021 showing the number of ensemble members that contributed daily precipitation, maximum temperature, and minimum temperature fields for the historical simulation and 8 different SSPs from ScenarioMIP. From Longmate et al, (In Revision)

Returning to Figure 2, the key step of this investigation, and the focus of this report, involves statistical and dynamical downscaling. The statistical downscaling focuses on the use of Localized Constructed Analogs (LOCA) (Pierce et al, 2014; Pierce et al, 2021). The dynamical downscaling of this research focuses on using the Weather Research and Forecasting (WRF) model (Powers et al, 2017; Skamarock et al, 2019).

Statistical Downscaling Overview

The starting point for the use of statistical downscaling is a recognition that there are myriad processes across a range of space and time scales that influence local temperature and precipitation patterns. The observational record naturally captures these processes at local scales. Furthermore, it may be challenging or even infeasible to represent these processes explicitly in the physical process models that comprise dynamical downscaling. The spatial and temporal scales at which some physical processes affect temperature and precipitation can be orders of magnitude higher than the highest resolution physical models, and even if computational resources were available to perform such calculations, the initial and boundary

conditions of process-resolving models would not be known, nor would there be any capability of independently assessing the structural uncertainty of such a model.

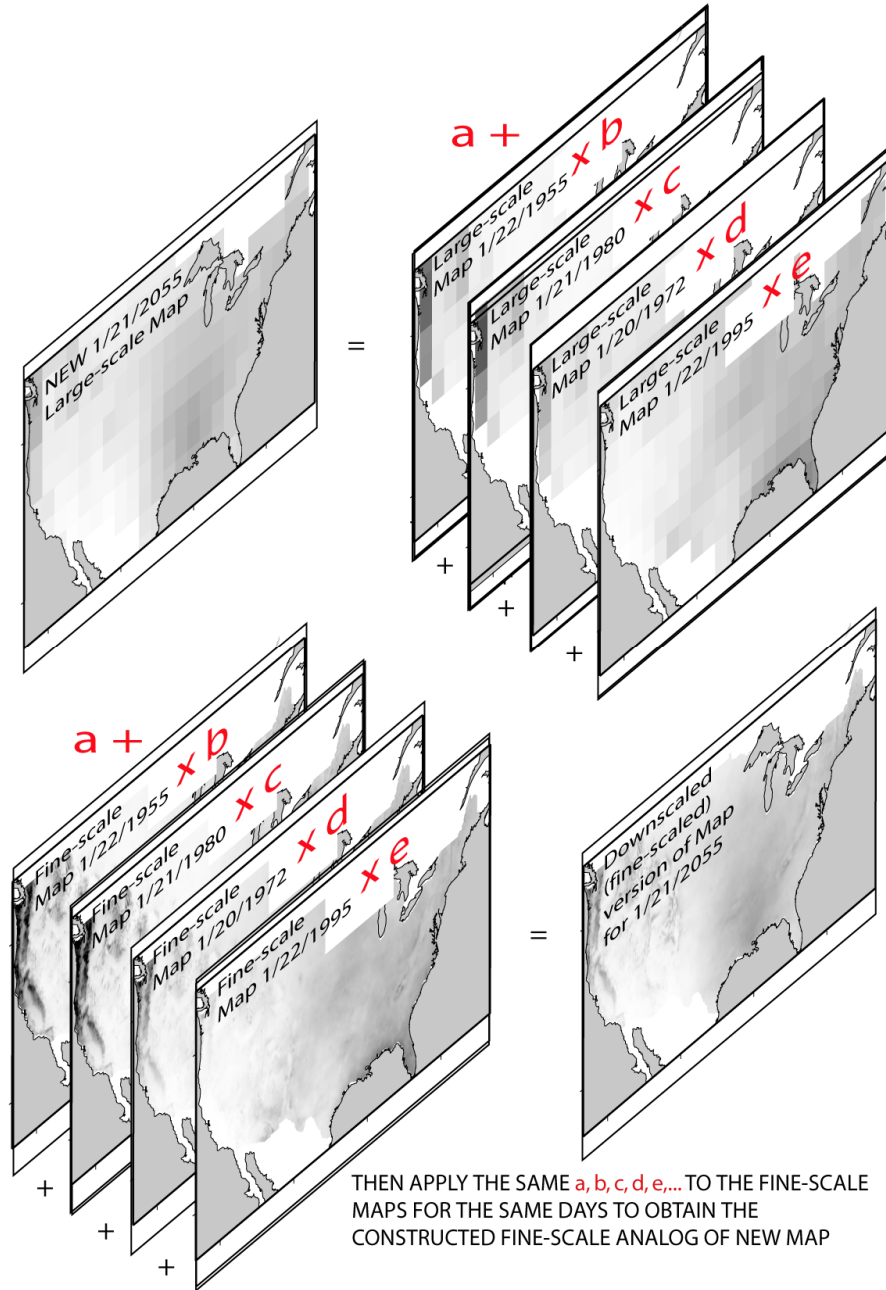


Figure 4: Pictorial diagram of statistical pattern matching that forms the basis of analog-day matching for statistical downscaling. From Hidalgo, H.G., Dettinger, M.D., and Cayan, D.R., (2008).

The LOCA method also recognizes the information contained in weather phenomena patterns and their spatial correlations on daily time-scales, since longer time-scales will convolve a number of processes while shorter time-scales will have aliased from diurnal cycles of many processes. Figure 4 depicts how LOCA uses a set of analog days over the historical record to

pattern-match the temperature and precipitation from an ESM in a historical simulation and leverages that pattern to project temperature and precipitation (or other downscaled variables of interest such as humidity and winds) from an ESM in a future projection.

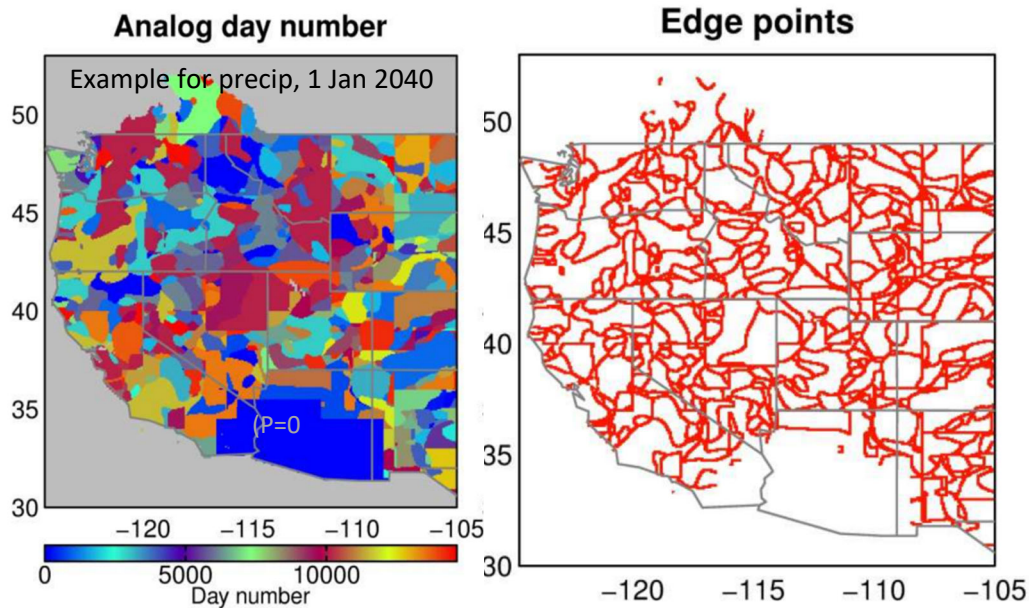


Figure 5: Analog-day matching for a sample CMIP5 model simulation showing the spatial variability in analog days used for statistical downscaling (left) and the boundaries between analog days for which spatial smoothing is required to avoid artifacts from spatially-disaggregation.

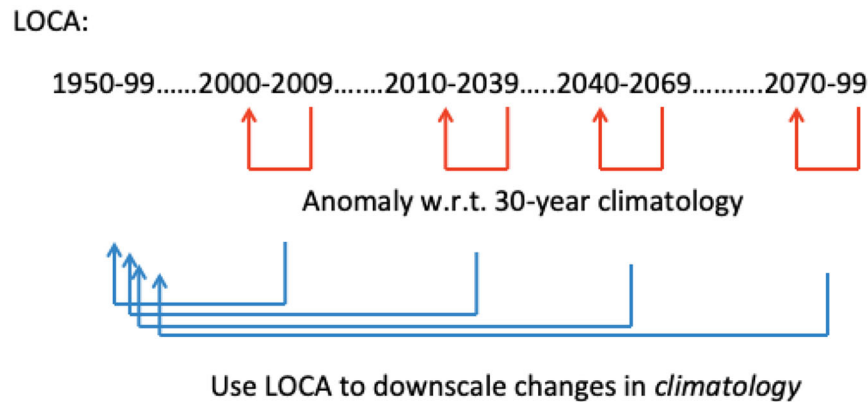


Figure 6: Challenge of using historical analog days for future projections of extreme events: the climatology of analog days in the past is fixed and is unable to develop projections beyond the range of historically-observed conditions.

However, the approach of pattern-matching with analog days does have the potential to introduce artifacts into downscaling. For example, the projection of an ESM may correspond to different analog days at different locations. Figure 5 (left panel) shows an example of analog day selection from LOCA, with Figure 5 (right panel) showing a tableau of different analog days for a given model projection day. This leads to potential discontinuities in downscaling fields that must be resolved through a spatial smoothing algorithm, as described in [Pierce et al, \(2014\)](#).

It also is challenged by the need to project extreme events in the future using a limited number of extreme events of the past that may not encompass extreme events of the future. Figure 6 shows a depiction of how projections of extreme events in the past develop downscaling solutions based on a small number of historical events.

To address this known issue, the statistical downscaling of this investigation focuses on producing a practical solution that can be, and has been, utilized by end-users. It does so through a utilization of the information on the relationships between weather phenomena and a wide range of Earth's climate states in the observational record. The investigation has developed important updates to LOCA by developing LOCA2 (Pierce et al, 2021). One of the innovations of LOCA2 is its time-of-day (TOD) adjustments (see Observational Product Analysis below). The second, and more pertinent, innovation addresses the question of extrapolating extreme precipitation in the future. LOCA used quantile mapping. The challenge with the use of QM is that the procedure cannot project more precipitation than the most extreme precipitation in the observational record. Figure 7 shows how the precipitation return period is capped in LOCA-downscaled CMIP5 models and produces an artifact in that the return value does not increase with a return period beyond 5 years. This artifact is addressed in LOCA2 through extreme value parameter fitting, as shown in the right panel of Figure 7, though it should be noted that the uncertainty in return value increases with return interval due to the sensitivity of the extreme value distribution to parameter choices and assumptions (Soukissian and Tsalis, 2015).

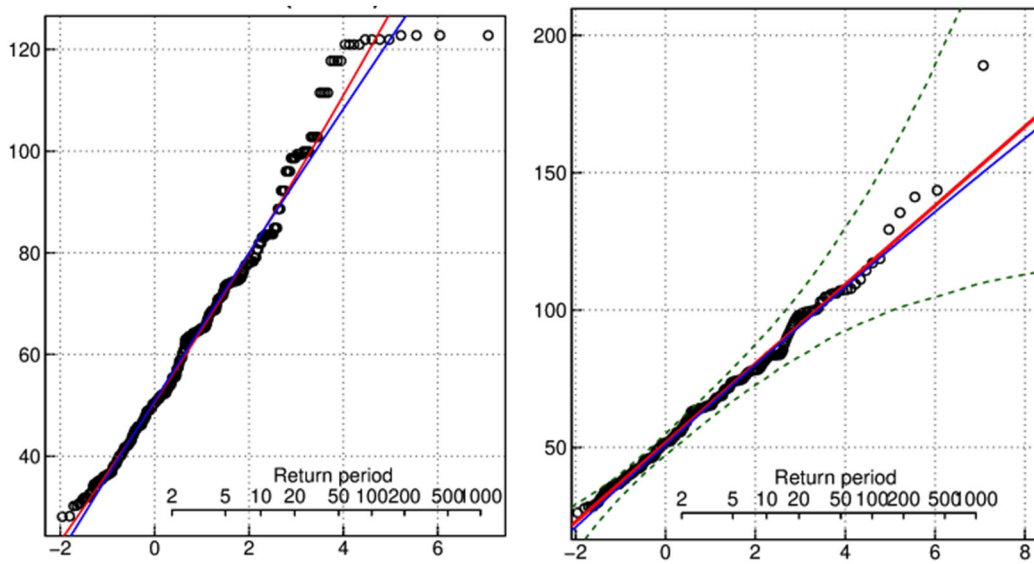


Figure 7: (Left panel) One-day precipitation extreme (y-axis) vs return period (x-axis) over Boston, Massachusetts in CMIP5-LOCA models. (Right panel) Same as left but showing results from LOCA2 applied to CMIP5 models. The lines depict least-squares fits to the return value vs. return period, depending on uncertainty in observations, while the dots correspond to CMIP5-LOCA or CMIP5-LOCA2 return values vs. return periods.

Figure 8 depicts a diagram of the process-chain of steps required to produce LOCA2-downscaling of CMIP6 models for daily precipitation, maximum temperature, and daily

temperature range (DTR) over the CONUS, as implemented on the NASA High-End Computing (HEC) facilities. The process-chain can be adapted to other variables where possible, such as surface humidity and winds. The process-chain starts by collecting a variable from a given emissions scenario and time-range from all CMIP6 parent models, crops those variables to North America, and creates data records on a standard calendar due to the fact that CMIP6 models have different calendaring systems (e.g., Sellar et al, 2020 notes that a year in the UKESM is 360 days).

It then performs Pass 1: one-level of downscaling to a 0.5×0.5 degree grid to account for the range of grid size (between 0.5 and 3 degrees) of CMIP6 parent models and performs a bias-correction based on the Livneh et al, (2015) (L15) dataset that has been regridded to a resolution of $0.5 \times 0.5^\circ$. The reason for this first downscaling is that there is a need to ensure that each LOCA2 downscaled CMIP6 model is augmented with the same amount of fine-scale information from statistical downscaling, which requires that they be processed onto the same initial grid. Otherwise, a parent model at coarse resolution would be augmented with far more information by LOCA2 than a parent model at fine resolution. After that, Pass 2 is performed to $1/16^{\text{th}}$ degree (6 km) resolution, wherein inputs are bias-corrected, downscaled, and then subjected to additional bias-correction.

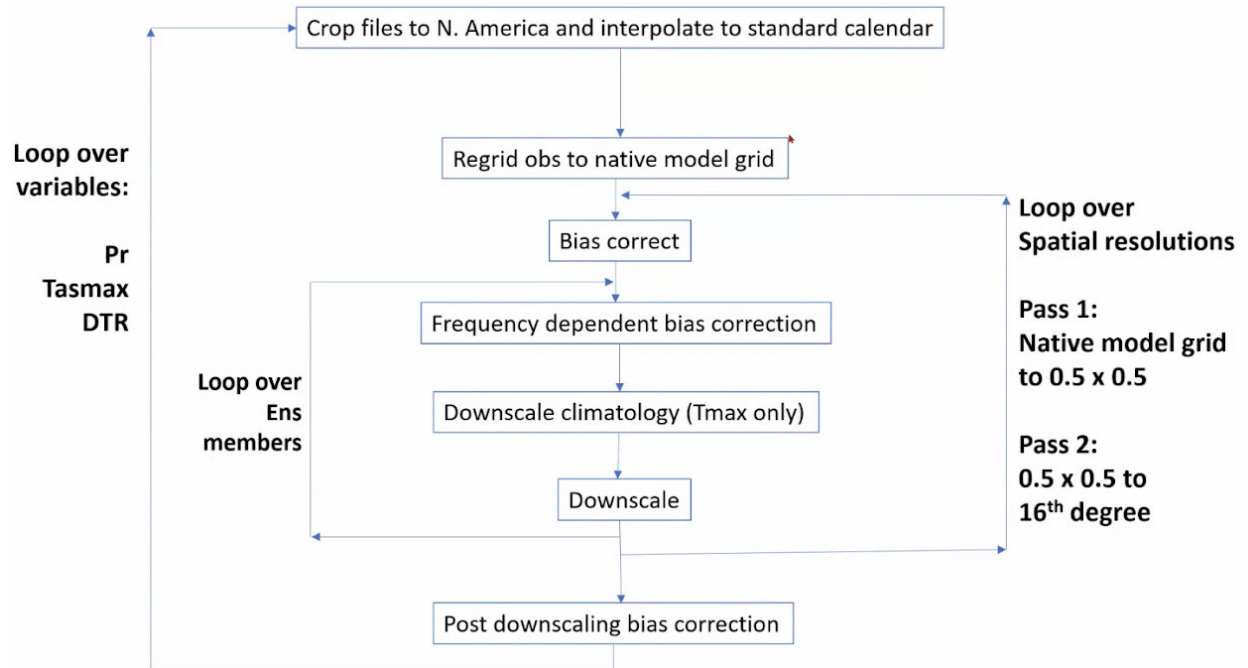


Figure 8: Diagram of LOCA2 process-chain for an ensemble of CMIP6 models.

LOCA has been revised to include updates to the training dataset and the treatment of extremes (described below). It has been implemented on NASA High-End Computing (HEC) facilities to produce ensembles of statistically-downscaled solutions for 27 ESMs in CMIP6 with between 1 and 10 initial-condition ensemble members per model for the historical, SSP245, SSP370, and SSP585 scenarios.

Table 2 highlights the CMIP6 models, their resolution, number of IC ensemble members and a reference to model simulations

CMIP6 Model	Resolution (Longitude X Latitude)	IC Ensemble Members	Reference
ACCESS-CM2	192 x 144	3	10.22033/ESGF/CMIP6.4271
ACCESS-ESM1-5	192 x 145	5	10.22033/ESGF/CMIP6.2288
AWI-CM-1-1-MR	384 x 192	5	10.22033/ESGF/CMIP6.2543
BCC-CSM2-MR	320 x 160	1	10.22033/ESGF/CMIP6.3016
CNRM-CM6-1	326 x 294	1	10.22033/ESGF/CMIP6.1375
CNRM-CM6-1-HR	720 x 360	1	10.22033/ESGF/CMIP6.1387
CNRM-ESM2-1	256 x 128	1	10.22033/ESGF/CMIP6.1391
CanESM5	128 x 64	7	10.22033/ESGF/CMIP6.1317
CESM2-LENS	288 x 192	20	10.22033/ESGF/CMIP6.7748
EC-Earth3	512 x 256	4	10.22033/ESGF/CMIP6.4700
EC-Earth3-Veg	512 x 256	5	10.22033/ESGF/CMIP6.727
FGOALS-g3	180 x 80	4	10.22033/ESGF/CMIP6.2050
GFDL-CM4	360 x 180	1	10.22033/ESGF/CMIP6.1402
GFDL-ESM4	360 x 180	1	10.22033/ESGF/CMIP6.1407
HadGEM3-GC31-LI	192 x 144	3	10.22033/ESGF/CMIP6.1901
HadGEM3-GC31-MM	432 x 324	2	10.22033/ESGF/CMIP6.1902
INM-CM4-8	180 x 120	1	10.22033/ESGF/CMIP6.5080
INM-CM5-0	180 x 120	5	10.22033/ESGF/CMIP6.5081
IPSL-CM6A-LR	144 x 143	9	10.22033/ESGF/CMIP6.1534
KACE-1-0-G	192 x 144	3	10.22033/ESGF/CMIP6.8350
MIROC6	256 x 128	5	10.22033/ESGF/CMIP6.898
MPI-ESM1-2-HR	384 x 192	10	10.22033/ESGF/CMIP6.762
MPI-ESM1-2-LR	192 x 96	10	10.22033/ESGF/CMIP6.6693
MRI-ESM2-0	192 x 96	1	10.22033/ESGF/CMIP6.621
NorESM2-LM	144 x 96	1	10.22033/ESGF/CMIP6.8036
NorESM2-MM	288 x 192	2	10.22033/ESGF/CMIP6.8036
TaiESM1	288 x 192	1	10.22033/ESGF/CMIP6.9702

Table 2: CMIP6 models that will be downscaled with LOCA2 and served publicly.

LOCA2 solutions from CMIP6 are available for community analysis at Globus with the following link: https://globus.nersc.gov/file-manager?origin_id=8b53db8a-d0e2-11ec-b95d-0f43df60473d.

Dynamical Downscaling Overview

In contrast to statistical methods, dynamical downscaling represents one of the most sophisticated approaches for producing comprehensive, high-resolution atmospheric fields over historical simulations or future projections. It achieves this sophistication by being the most

physically complete downscaling approach: it ingests coarse-resolution information from a climate model, consisting of roughly a dozen variables that the model produces, and then uses the primitive equations, to create what is effectively a regional climate model's (RCM's) dynamical core (the term "effectively" refers to the fact that the dynamical downscaling model does not need to be explicitly compatible with the parent model, as an RCM generally does) that outputs a solution that represents the local meteorology. By explicitly including the representation of primitive equations, dynamical downscaling has, in principle, the ability to make projections that account for nonstationary effects because physical models can predict meteorological conditions of the future even if such conditions have not been observed in the historical record. The downscaling RCM is generally held to mass, energy, and water-budget balance conditions. The higher-resolution topographic variables used by the RCM (e.g., land/sea classification, elevation, vegetation fraction, etc.) also allow for the simulation of a more realistic surface-atmospheric coupling. ^[L]_[SEP]

A major limitation inherent to dynamical downscaling is that, although computing capabilities are improving, it can be far too expensive to develop dynamical downscaling solutions over climate time scales (Wrzesien & Pavelsky, 2020). ^[L]_[SEP] Additionally, it takes a specific set of variables, available at sub-daily intervals, to drive RCMs; these data are not as widely available as are daily and monthly data, particularly for GCMs. A further limitation of dynamical downscaling is that regional climate modelers have wide latitude in choosing initialization procedures, RCM physics, and grid options, and this creates structural uncertainty and leads to variations in the regional solutions. Such choices are not easily tested on climate time scales due to the aforementioned computational requirements, making it difficult to ascertain their effect on model realism. In point of fact, even though this project is expending millions of CPU hours on world-leading supercomputing platforms (as shown <https://dept.atmos.ucla.edu/alexhall/downscaling-cmip6> and <https://registry.opendata.aws/wrf-cmip6/>) to develop dynamical downscaling solutions, it is able to downscale 6.1% (20 vs 329) of the centennial-length GCM simulations that are being statistically-downscaled with LOCA2. Simply put, the range of GCM simulations in CMIP6 cannot be sampled exhaustively with dynamical downscaling. The ambitious effort being undertaken by this project still must be heuristically assessed to determine how sufficient it is in capturing the range of model projections across the CONUS.

Nevertheless, this investigation has stood up a comprehensive dynamical downscaling capability based on the Weather Research and Forecasting Model (WRF) (Rahimi et al, 2022). This capability has been developed and tested extensively over the historical period of observation in order to demonstrate its utility as a tool for translating large-scale dynamic and thermodynamic changes in the Earth's atmosphere to future projections at the local scale. The WRF configuration has an inner domain with 3-km resolution over California with a middle domain at 9 km resolution over the Western United States (WUS), as shown in Figure 9.

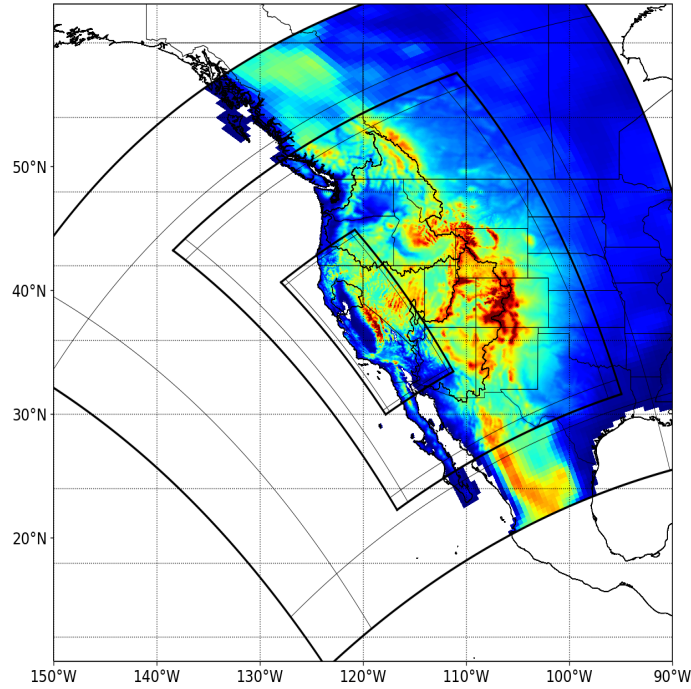


Figure 9: Simulation domains of Western United States (WUS) dynamical downscaling with WRF for RC19-1391. The inner-domain, at 3 km resolution, covers California, while the middle-domain, at 9 km resolution, covers the WUS. From Rahimi et al, (2022).

The process for determining an optimal configuration of the WRF model for dynamical downscaling over the WUS was highly non-trivial. WRF has many different parameterization configurations that are appropriate and well-suited for some applications of that model and not for others (Lavin-Gullon et al, 2021). This investigation tested the major uncertainties in dynamical downscaling configuration: WRF forcing and parameterizations. Figure 10 shows the significant impacts of different choices of forcings and parameterizations on the skill of the WRF configuration over the historical period of performance.

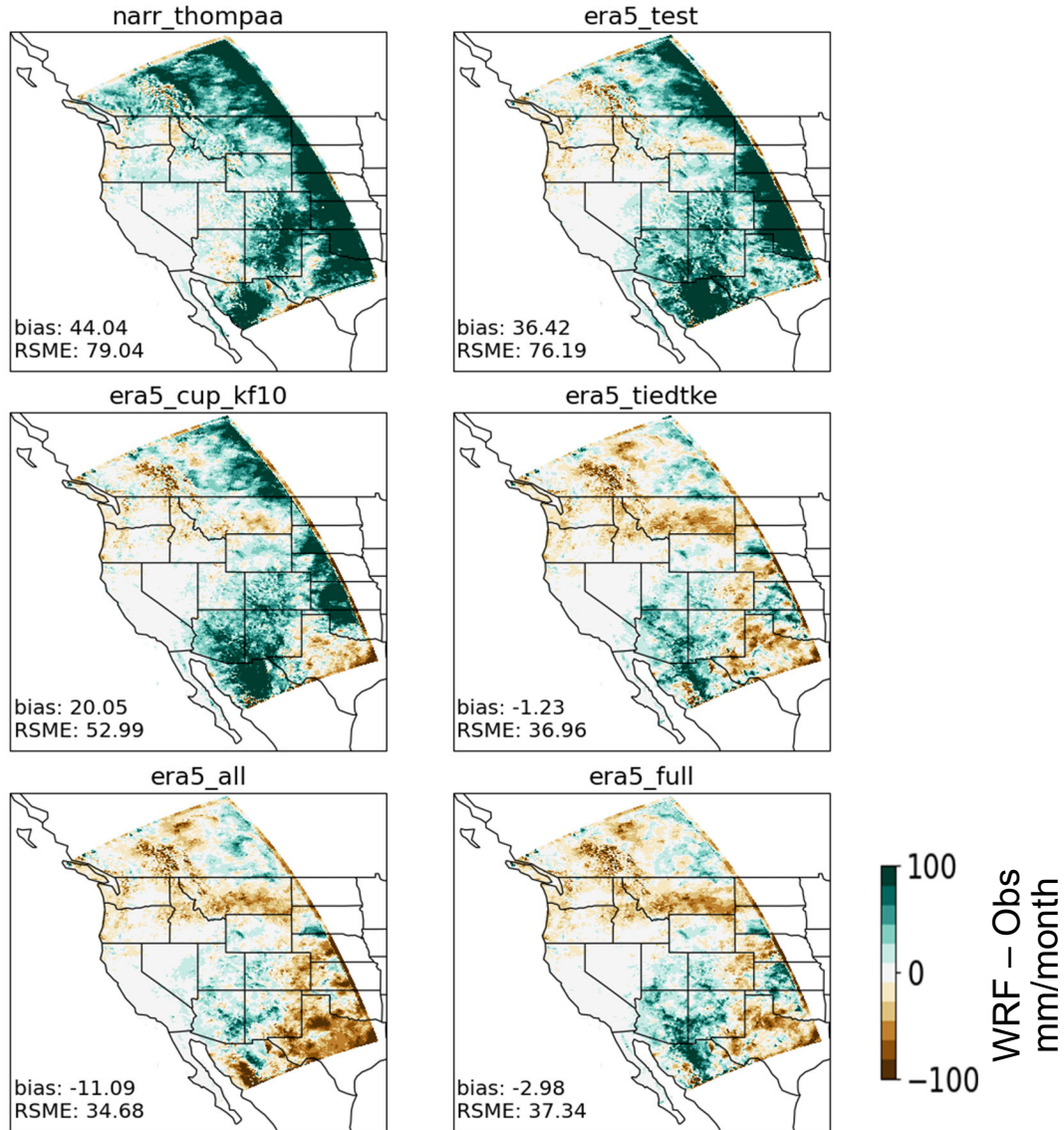


Figure 10: Maps of 1980-2009 difference between several WRF configurations and the PRISM observational dataset over the WUS. The top row shows the effect of WRF large-scale forcing by comparing NARR and ERA5 forcings. The middle row shows the effect of different cumulus parameterizations (Tiedke vs Kain), and the lower row shows the effect of different land-surface configurations (Noah-MP (all) vs Noah-MP with `opt_rad=1` (full)).

Parameterization	Initial Experiment	Final experiment
Microphysics	P3	P3
Cumulus	Kain (2004)	Tiedke (1989); Zhang et al, 2011
Land surface	Noah-MP	Noah-MP with <code>opt_rad=1</code>
Surface layer	Revised MM5	Revised MM5
Planetary Boundary Layer	YSU	YSU
Nudging on coarsest grid	Spectral: u, v, T, q	Spectra: u, v, and T
Downscale to 45-km grid first?	No	Yes

Driving reanalysis	NARR	ERA5
--------------------	------	------

Table 3: Summary table of WRF initial and optimized (final) configurations used for dynamical downscaling of WUS in this investigation.

The results of the engineering activities associated with developing an optimized WRF configuration are highlighted in summary statistics of the performance of WRF relative to two observational products over the historical record. Figure 11 shows PRISM and Daymet (Thornton et al, 2014). These show a very low bias and RMSE for temperature and precipitation anomalies relative to PRISM and Daymet.

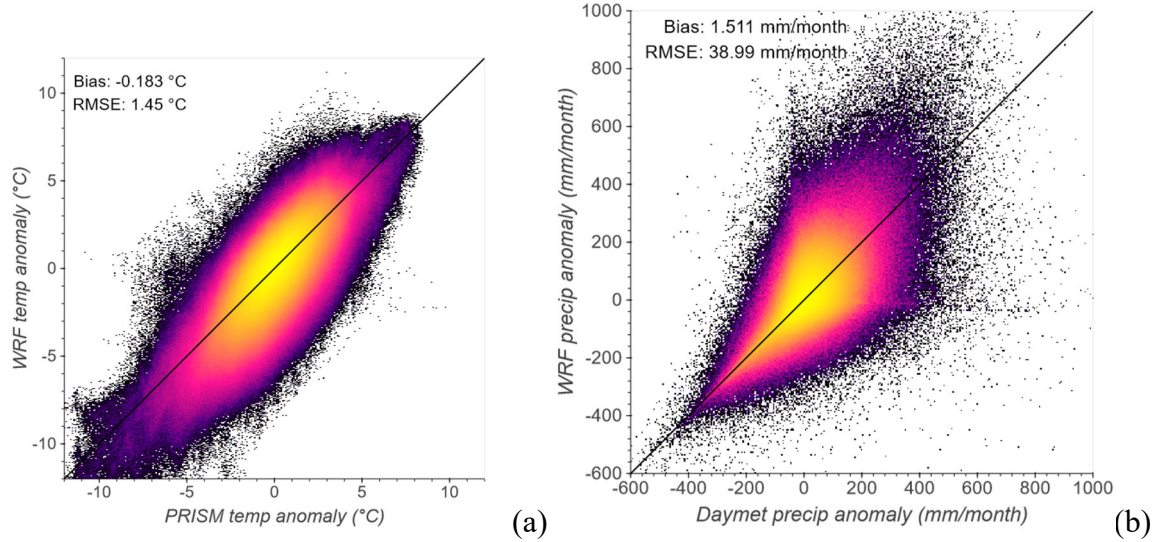


Figure 11: Scatterplot/heat-map of performance of WRF WUS configuration from 1979-2010 relative to PRISM and Daymet observational datasets.

This optimized WRF configuration has been described in Rahimi et al (2022) as a high-throughput dynamical downscaling capability for CMIP6 whereby a configuration of WRF receives initial and boundary conditions for a given CMIP6 model. To date, this investigation has used this WRF configuration to produce dynamically-downscaled simulation of the WUS from 1980-2100 for SSP370 for the following models: CNRM-ESM2-1, CESM2, and MPI-ESM1-2-LR.

Nonstationary Analysis of Downscaling Solutions

This investigation focuses on developing statistical and dynamical downscaling solutions over the CONUS, but, as described above, these solutions contain myriad different assumptions: one is based on historical pattern-matching, while the other is focused on solving the primitive equations. The divergence of the two solutions in space and time gives a strong indication of the potential importance of nonstationarity effects in influencing downscaled projections.

There are strengths and weaknesses to each of the downscaling solutions developed here. Because dynamical downscaling is based off of solving primitive equations, such solutions can

contain realistic nonstationary effects that may occur in the 21st Century but have yet to be observed in the historical record. The expectation, therefore, is that, *a priori*, dynamical downscaling is more likely to capture nonstationary effects than statistical downscaling.

At the same time, it should be noted that there are a wide range of parameterizations with any dynamical downscaling approach which are approximations of actual physical processes and are created based on present-day conditions. The parameterizations of the components of a dynamical downscaling model and the interactions of those components are calibrated against historical observations. Therefore, dynamical downscaling solutions may have subtle stationarity assumptions that are built into the model parameterizations and their interactions through the calibration process. For example, dynamical downscaling contains a microphysical scheme to describe how local atmospheric winds, humidity, temperature, and aerosols create local precipitation. The dynamical downscaling model could be run at resolutions down to tens of meters (through coupling to Large-Eddy Simulations), but even then, the actual formation of precipitation occurs on spatial scales of millimeters and must be parameterized. Microphysical processes are underrepresented in any of those microphysical schemes and they vary in space and time to such an extent that in the future, a different microphysical scheme could easily be more appropriate than the scheme used in present-day or historical dynamical downscaling. Because precipitation amount, phase, and timing are strong functions of the microphysics scheme, there is an implicit stationarity assumption in dynamically-downscaled precipitation associated with microphysical scheme choice.

At the same time, statistical downscaling is built off of historical observations. While it is not correct to say that statistical approaches cannot model nonstationary effects, statistical methods do model nonstationary effects in different ways. For example, in LOCA2, extreme precipitation is actually modeled by a Generalized Extreme Value (GEV) distribution, with parameters estimated from historical observations. GEV distributions can predict conditions that do not have historical analogues, but the parameters of the distribution are stationary, and so if the underlying physical processes that are represented in the GEV parameters change, then the GEV distribution will not be able to capture nonstationary effects and therefore LOCA2 will not be able to handle nonstationary effects in extreme precipitation.

So what is the way out of this seemingly unresolvable conundrum? It is important to note that both downscaling approaches have their strengths and weaknesses, and it is incumbent on the community to combine the strengths of these approaches while minimizing their weaknesses. The most straightforward starting point for combining strengths is simply to compare statistical and dynamical solutions for models over the same experiment. Agreement between the two solutions is a strong indicator that nonstationary effects are not likely to be significant, since the two methods arrive at solutions in drastically different ways. Where they do diverge, systematic downscaling calculations and analyses are required and that work is location and time specific.

The following is a case-study of an approach to identify such divergence and what can be learned upon further investigation about the sources of the divergence and the role of nonstationary effects in influencing the downscaling solution. [Walton et al, \(2020\)](#) compares three downscaled datasets over the Sierra Nevada to examine differences based on training data and differences based on model relationships. The datasets are (1) dynamically downscaled results from the

Weather Research and Forecast (WRF) model, a state-of-the-art RCM; (2) statistically downscaled results using LOCA, a state-of-the-art statistical model, trained on historical WRF output, which is termed “LOCA-WRF”; and (3) statistically downscaled results using LOCA, this time trained on the Livneh station-based gridded observational dataset (Livneh et al, 2013), which is termed “LOCA-Livneh”. By comparing WRF and LOCA-WRF, Walton et al (2020) removes the effect of differences between the training dataset and WRF climatology, and isolates disagreements due to different downscaling procedures between WRF and LOCA. On the other hand, comparing LOCA-WRF and LOCA-Livneh isolates disagreements due to the choice of the training dataset’s climatology (WRF versus Livneh) in the statistical model.

Figures 12 and 13 shows significant differences between 3 different statistically-downscaled solutions (d, e, and f) and the WRF solution for the change in April surface temperatures, with the CNRM-CM5 model (GCM) field approximating, at coarse resolution, the WRF solution is manifestly different, especially with respect to elevational change in temperature in April. Similarly, the patterns of change in downscaled solutions between WRF and the statistically downscaled solutions highlight southeastern California as an area of downscaling solution divergence.

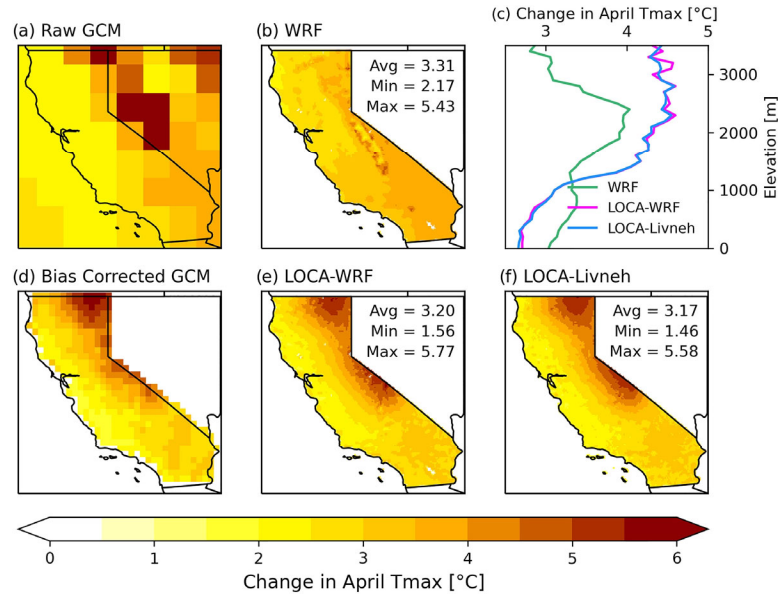


Figure 12: End-of-century (2091–2100 average minus 1991–2000 average) changes in April maximum temperature according to (a) the raw GCM (CNRM-CM5), (b) WRF, (d) bias-corrected GCM, (e) LOCA-WRF, and (f) LOCA-Livneh. The statewide average, minimum, and maximum for each downscaled data set are reported in the upper right margin of those panels. Unit for each panel is degrees Celsius. Panel (c) compares change in April maximum temperature versus elevation for each grid cell within the WRF (green line), LOCA-WRF (magenta line), and LOCA-Livneh (blue line) projections. From Walton et al (2020).

Walton et al (2020) performs a set of closer investigations as to the causes of this downscaling solution divergence and found that they can be linked to errors in GCM snowpack (for temperature) and errors in the contrast in temperature and radiation between the ocean and the North American continental interior (for precipitation). In terms of snowpack, snow-albedo feedbacks significantly impact temperature trajectories, so if a GCM projects too much

precipitation amounts in mountains, it will also project too much snowpack. Consequently, statistically-downscaled solutions will be forced by that pattern to produce too-cold temperatures. This analysis shows that, in this case, the source of the divergence in statistical and dynamical solutions is the impact that precipitation has on the snowpack, which in turn impacts how much a model can warm in the future. To address errors caused by the overestimation of regional precipitation by parent models, dynamical downscaling can be used, or, alternatively, statistical downscaling can be used so long as the parent model is bias-corrected first.

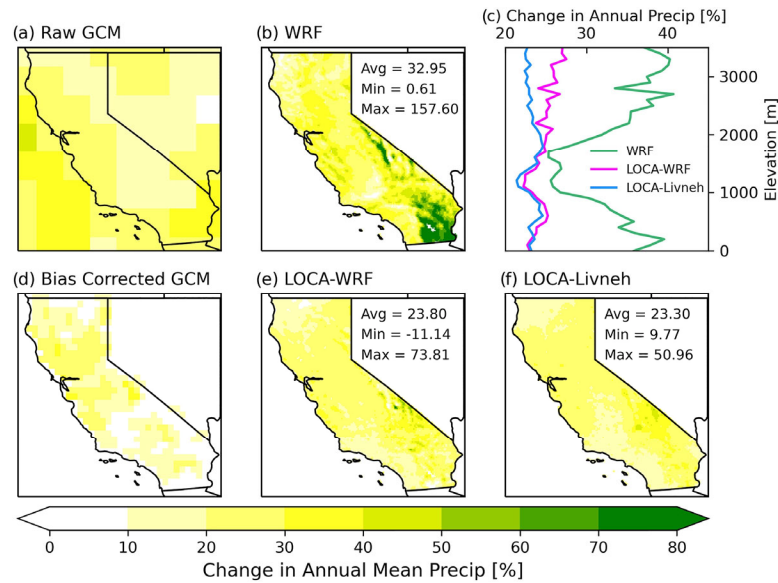


Figure 13: End-of-century (2091–2100 average minus 1991–2000 average) changes in annual mean precipitation according to (a) the raw GCM (CNRM-CM5), (b) WRF, (d) bias-corrected GCM, (e) LOCA-WRF, and (f) LOCA-Livneh. The statewide average, minimum, and maximum for the downscaled products are reported in the upper right margin. Unit for each panel is percent. Panel (c) compares the change in annual mean precipitation versus elevation for each grid cell within the WRF (green line), LOCA-WRF (magenta line), and LOCA-Livneh (blue line) projections. From Walton et al (2020).

Meanwhile, Figure 13 shows very different annually-averaged precipitation patterns over California between WRF and all the other solutions, particularly in southeastern California. There, precipitation is the result of a ocean-land temperature contrast drives the North American Monsoon (NAM). The NAM is responsible for nearly all precipitation in southeastern California, so if a GCM has too much ocean-land contrast, then any statistically-downscaled solution will produce too much precipitation in that area. Again, accurate projections of precipitation in southeastern California require an accurate ocean-land temperature contrast. That can be achieved either with dynamical downscaling or, alternatively, statistical downscaling can be used so long as the parent model is bias-corrected first.

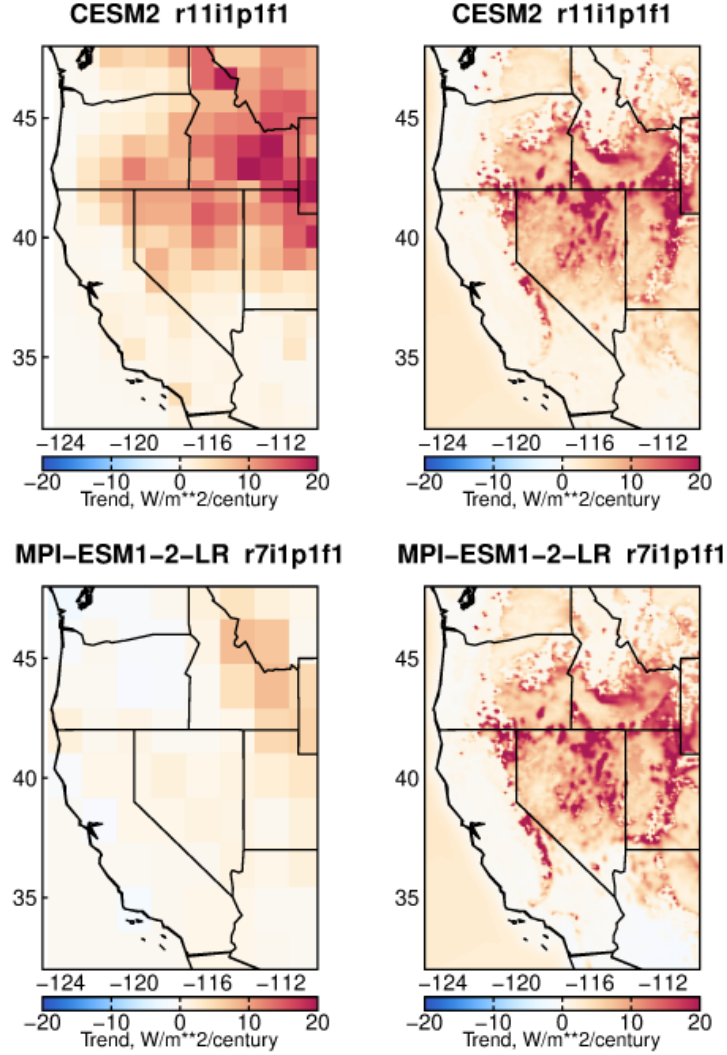


Figure 14: Trend from SSP370 for the 21st Century (2014-2099) in net surface shortwave radiative energy (down-up) in (upper-left) raw CESM2 simulations, (upper-right) bias-corrected WRF-downscaled CESM2 simulations, (lower-left) raw MPI-ESM1-2-LR simulations, and (lower-right) bias-corrected WRF-downscaled MPI-ESM1-2-LR simulations.

The efforts by Walton et al, (2020) indicate how and where this can work in California: they ran a series of different downscaling solutions using LOCA trained with different historical datasets and also datasets constructed for future climatic conditions using dynamical downscaling. They used these to isolate the physical processes that were leading to the divergence between the statistical and dynamical methods and identified how precipitation, snowpack, and Gulf of California circulation are prerequisites for accurate downscaling solutions of temperature in the mountains of California and precipitation in southeastern California. That exacting approach, which varies by location, is a path forward to develop robust downscaling solutions.

Bias-correction is another key aspect to downscaling divergence analysis for understanding nonstationary effects. While this project has shown that bias correction is important, especially

for statistical downscaling, it too must be used carefully for developing downscaled projections. Bias correction is a process by which GCM biases relative to a reference dataset are removed. Many bias correction methods have been developed for use in statistical downscaling (Ehret et al, 2012). Bias correction in this context is often applied to surface air temperature and precipitation, or other variables of interest to end-users and stakeholders, making use of observationally-based datasets. While there is a robust, ongoing scientific debate about the utility of bias-correction, inarguably, additional information is being introduced to the downscaling solution when bias-correction is applied. There are open questions as to the value of that information, its consistency with the dataset being bias-corrected, and overall what is learned when adjustments are made to datasets.

To address some of these questions, this project has recently looked into the impacts of bias-correction on dynamical downscaling trends. The background here is that while statistical downscaling solutions require bias correction, dynamical solutions are also impacted by biased initial and boundary conditions from the parent GCM. Specifically, the investigation is exploring whether bias-correcting the parent model impacts precipitation and temperature trends with dynamical downscaling. Where it does, this investigation is exploring when and where significant differences arise between bias-corrected and non-bias-corrected dynamical downscaling in order to use this as a proxy to assess whether there are stationarity assumptions in bias-correction itself (i.e., does it have a “shelf-life”?)

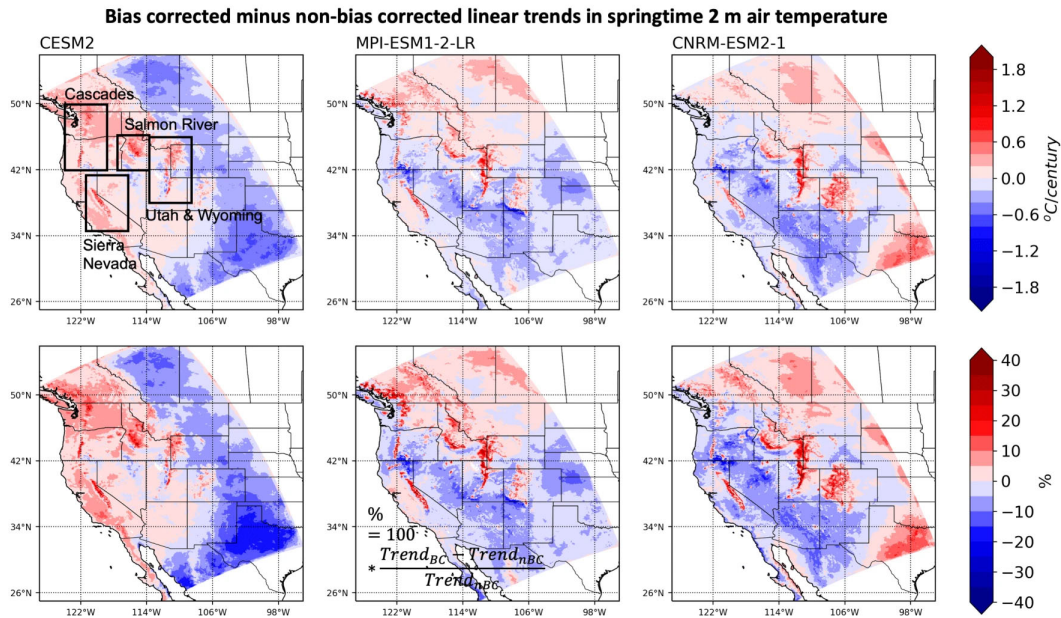


Figure 15: Differences in springtime (March-April-May average) 2-meter air temperature trends of bias-corrected and non-bias-corrected WRF dynamical downscaling for CESM2, MPI-ESM1-2-LR, and CNRM-ESM2-1 models for 1980-2100. Top row shows differences in degrees/Century and bottom row shows percent difference in trends.

The CESM2, MPI-ESM1-2-LR, and CNRM-ESM2-1 models have been dynamically downscaled with WRF and integrated over the WUS from 1975-2100 for two separate

integrations: with bias-correction of the models prior to their fields being ingested into WRF through the application of the method of [Bruyère et al., \(2014\)](#) and without such bias-correction, as shown in Figure 14. Many aspects of bias exist in these parent models, so the following variables were bias-corrected: three-dimensional temperature, wind and humidity fields, sea-surface temperature, soil moisture and temperature, surface pressure.

The results in Figure 15 show that there are distinct spatial patterns of the influence of bias correction on 2-meter air temperature trends. Four high-elevation locations in the mountain west are highlighted because the higher elevations show increased temperature trends in bias-corrected fields relative to non bias-corrected fields. As has been found in [Walton et al, 2020](#), the elevational dependence of the impact of bias-correction on temperature trends is suggestive of a role of snow-albedo.

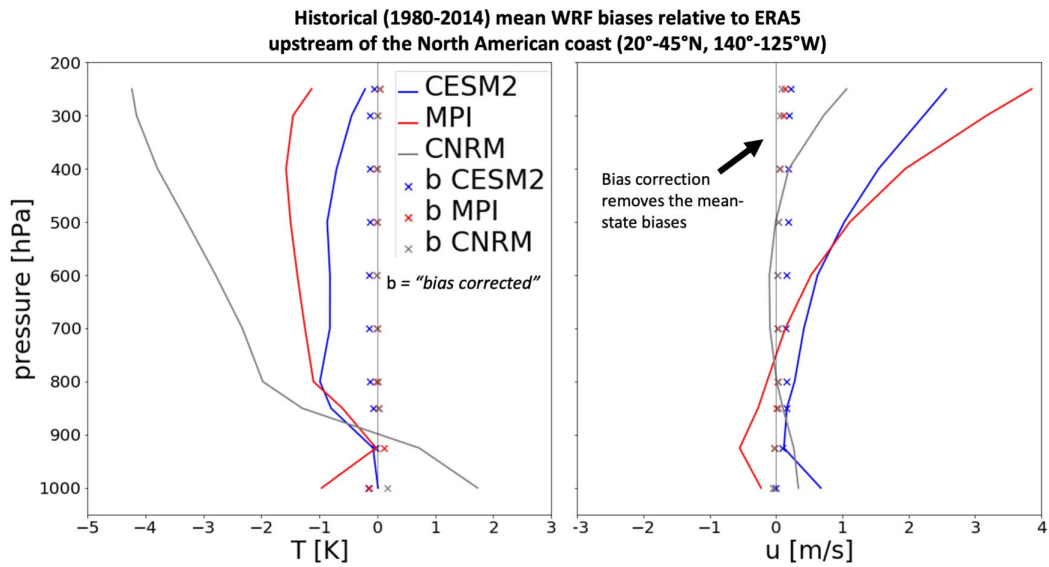


Figure 16: Temperature and wind-speed profile mean biases in 3 dynamically-downscaled models relative to ERA5 over 1980-2014 upstream of the North American west coast.

A deeper analysis of the models shown in Figure 16 reveals the causes for the different parent model trends. By looking at the biases in upstream thermodynamic temperature structure and vertical winds in the historical model simulations relative to ERA5, it is apparent that the static stability and vertical shear biases in the MPI-ESM1-LR model lead that model to produce too much precipitation over the WUS. This upstream bias in MPI-ESM1-LR has been found to be consistent during the historical simulations and is suggestive of persistent model errors upstream of the CONUS and, by extension, a need for bias-correction in dynamical downscaling.

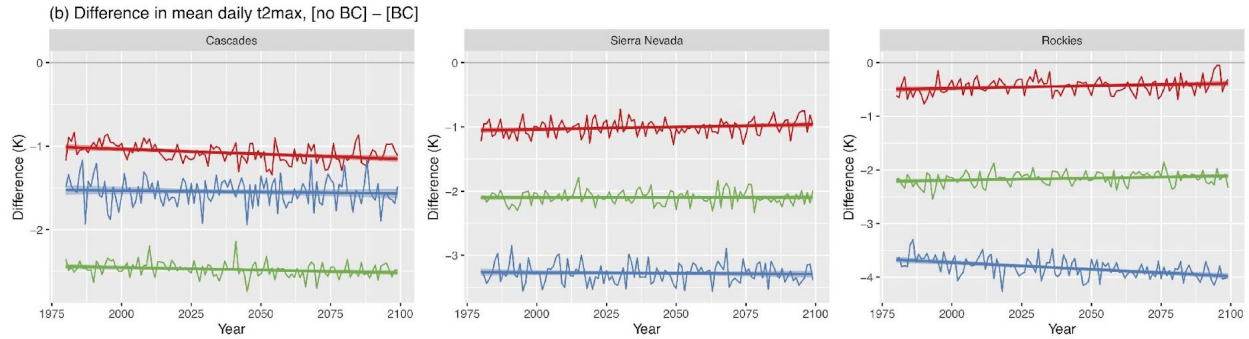


Figure 17: Differences in daily 2-meter maximum air temperature over the Cascades, Sierra Nevada, and Rockies for non-bias-corrected and bias-corrected WRF simulations from 1975-2100. Red is CESM2, blue is MPI-ESM1-2-LR, and green is CNRM-ESM2-1.

This point is emphasized when looking at whether there are time-dependent differences in the trends between bias-corrected and non-bias-corrected WRF simulations. Figure 17 shows little-to-no time-dependent differences in trends, suggesting that the bias, and likely the underlying causes of the bias, observed in the historical simulations of the model are carried forward with future simulations, and that a bias-correction based on historical model performance in these high-elevation locations of the WUS should be applied to downscaling solutions. Furthermore, these results suggest that bias-correction has a long “shelf-life” and is appropriate for developing realistic dynamical downscaling over the WUS.

In general, divergence of controlled dynamical and/or statistical downscaling experiments is a red-flag for the scientific community that there are substantive disagreements between statistical methods and atmospheric modeling. It is also indicative that the scientific community should look into the underlying reasons for these disagreements and the practitioner community should recognize that there is greater potential for an underestimation of risk due to changing temperature and precipitation distributions in the areas of downscaling solution divergence.

In order to investigate when and where downscaling solutions are diverging, this investigation is focusing on the CMIP6 archive to develop a wide-range of statistical and dynamical downscaling solutions. Whether or not the dynamical method is more realistic than a statistical method and captures nonstationary effects of the future requires further process-based investigation.

Parsimoniously Downscaling CMIP6

As shown in Figure 3, the CMIP6 archive contains numerous separate contributions from different modeling centers, and it also contains different ensemble members of each model produced through changing initial conditions (IC) or perturbed physics. These intra-model ensembles enable assessments of a given model’s internal variability, with IC ensemble members representing, by far, the most common ensemble in CMIP6 (as of October 11, 2021, only a single model had contributed a perturbed-physics ensemble to ESGF). A recent paper by Mankin et al. (2020) showed the importance of IC ensemble members to decision-making for

hydroclimate. However, there is limited guidance in the scientific literature for navigating a multi-model ensemble of ESMs where some models contribute multiple ensemble members. Unfortunately, the relationships between intra-model uncertainty and downscaled hydroclimate projections have been grossly under-investigated.

One major previous effort to downscale for the CMIP5 archive with LOCA (Pierce et al, 2014; Kunkel et al, 2016) used a convention of the first IC ensemble member to produce one downscaled climate model per emissions scenario, which assumed that each IC ensemble member was sufficiently similar to the others for downscaling purposes. The sufficiency of selecting a single ensemble member for developing downscaled solutions and sampling the multi-ensemble has not, to date, been established. This investigation has explored how best to sample these simulations, and whether that sampling should expend resources to include different IC ensemble members or if the sampling should favor a greater range of models.

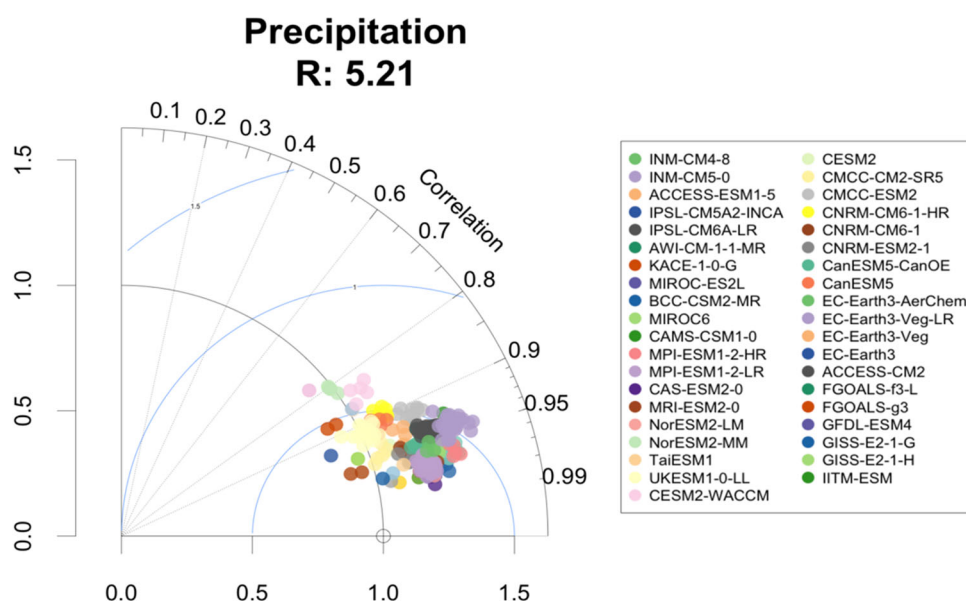


Figure 18: Taylor Diagram of the correlation and standard deviation of 31 CMIP6 models and their initial-condition ensemble members of historical simulations relative to the Livneh et al, 2015 hydroclimatological dataset for daily precipitation. From Longmate et al (In Revision).

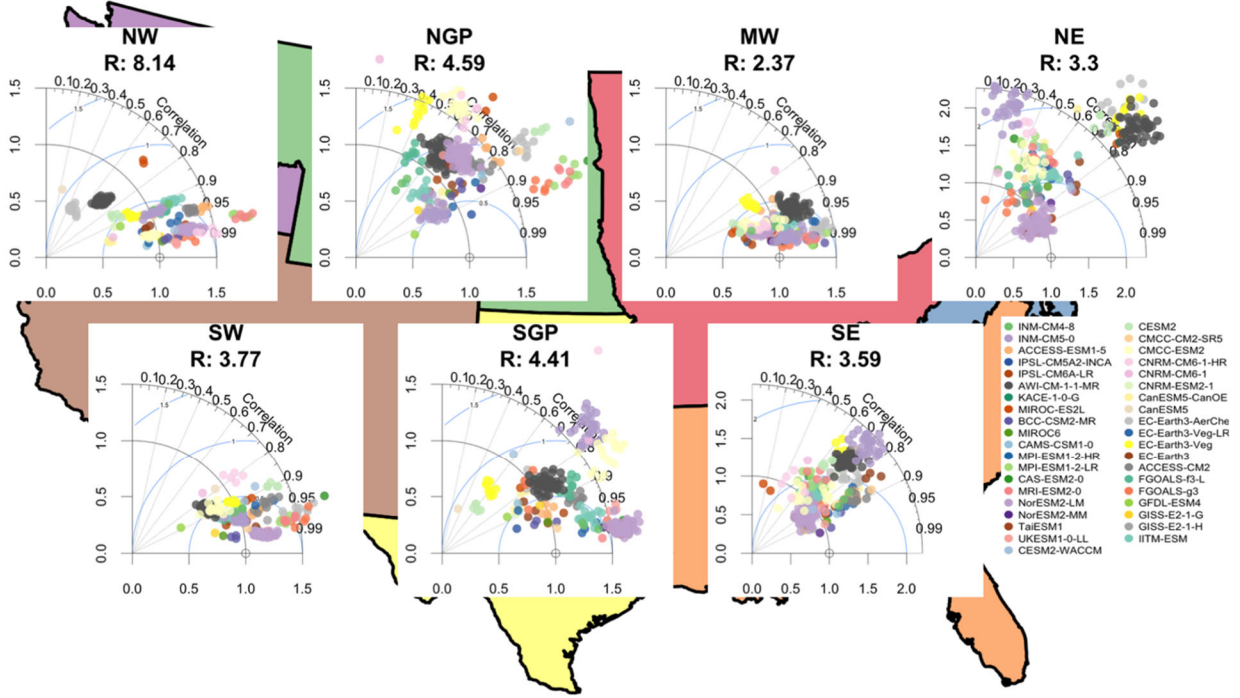


Figure 19: Taylor Diagram of the correlation and standard deviation of 31 CMIP6 models and their initial-condition ensemble members of historical simulations relative to the Livneh et al, 2015 hydroclimatological dataset for three sub-regions of the CONUS: the Western US, the Central US and the Eastern US. From Longmate et al, (In Revision)

The mean-centered statistics of a Taylor diagram (Taylor, 2001) are designed to concisely summarize the degree of correspondence between two fields or “patterns,” here comparing each ensemble member with the reference data set (either the multi-model ensemble average or the L15 product).

The statistics of interest are the standard deviation of each ensemble member $\{\sigma_{ij}; i = 1 \dots N; j = 1 \dots n_i\}$ (where N is the total number of models and n_i is the number of ensemble members for model i) and the reference data set σ_r as well as the correlation coefficient between each ensemble member and the reference data set, denoted $\{\rho_{ij}; i = 1 \dots N; j = 1 \dots n_i\}$. Both quantities are calculated across grid cells. These statistics can be further aggregated into the so-called centered pattern root mean square (RMS) difference D for each ensemble member, denoted $\{D_{ij}; i = 1 \dots N; j = 1 \dots n_i\}$, which can be written in terms of the standard deviations and correlation coefficient as:

$$D_{ij} = \sqrt{\sigma_{ij}^2 + \sigma_r^2 - 2\sigma_{ij}\sigma_r\rho_{ij}} \quad (1)$$

Note that as D_{ij} approaches zero as the two fields or patterns become more alike. Furthermore, it should be noted that the RMS differences do not account for overall mean differences in the two fields. In the following, RMS differences generically refer to a specific variable (precipitation or temperature) in a specific time period (historical or future) for a specific region (CONUS or one

of the sub-regions). The Taylor diagram visualizes these three statistics (centered pattern RMS difference, standard deviation, and correlation coefficient) in a single plot. The reference data set used is either the multi-model ensemble average or the L15 product. The model-ensemble average is normalized and ensemble members are normalized by the model-ensemble average.

As shown in Figures 18 and 19, across CONUS and for different sub-regions, models and ensemble members are quite similar to each other for historical simulations. Each cluster of similarly-colored points is closer to itself than to points of different colors, indicating that IC ensemble members of a given model exhibit similar statistics to each other, while IC ensemble members of another model generally are substantially different from each other. This result is suggestive that downscaling a single model is a judicious choice for parsimoniously downscaling the CMIP6 archive with limited computational and personnel resources.

Expanding on this analysis and to quantify whether choosing a single model is a judicious choice both for historical simulations and SSPs, this investigation has undertaken an additional analysis beyond the use of Taylor diagrams. While Taylor diagrams are helpful for visualizing a large amount of information about IC ensemble behavior between models and within models, the units of distance between points on are not uniform or easily interpretable. Even when restricting scope of the investigation to monthly climatologies over CONUS and various sub-regions, the scope of analysis across the large amount of information in the CMIP6 multi-model ensemble, and across different variables, scenarios, spatial extents, and time periods is too large, and necessitates distilling down the mean-centered statistics of a Taylor diagram into a single value to describe differences in between- and within-model variability.

In order to assess whether a random selection of model ensemble members will provide an unbiased multi-model ensemble for downscaling temperature and precipitation in a particular region, we can use a standard statistical approach called random effects modeling that quantifies the magnitude of between-group variability relative to the within-group variability. In this case, “group” refers to a climate model, with the ensemble members comprising the items in each group; the quantity of interest is the centered RMS differences and its variability across and within models. The standard setup specifies that the ensemble members from each model are a random sample of all possible IC ensembles, where each model has a model-specific mean centered RMS difference, say $\bar{D}_j$ and some variance τ^2 that does not depend on the model, which will indicate the “within-model” variability. For convenience, a Gaussian process (GP) distribution is used.

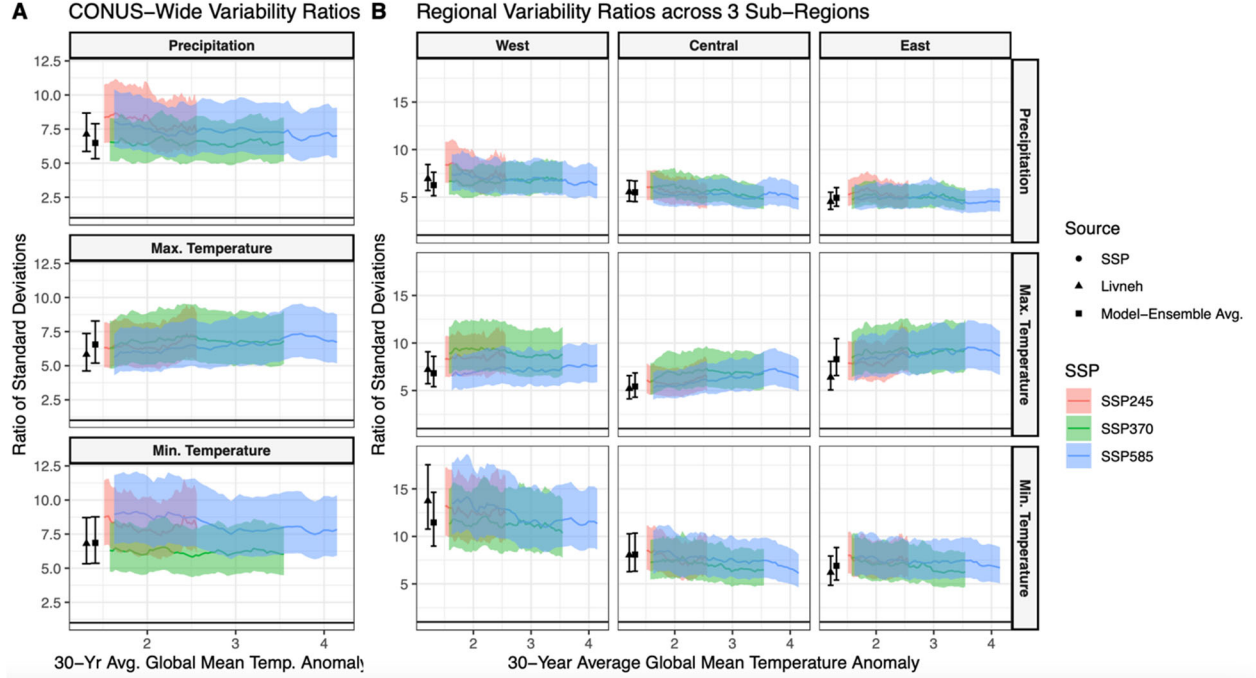


Figure 20: The variability ratio of between-model to within-model standard deviations across rolling 30-year windows from 2015 to 2100, plotted versus the average change in global mean temperature during that 30-year period for shared socioeconomic pathway (SSP) 245 (red), 370 (green), and 585 (blue). Panel A shows CONUS-wide variability ratios for precipitation (top), maximum temperature (center), and minimum temperature (bottom), while panel B shows regional variability ratios across three regions (West, Central, and East) for the same three variables. The corresponding variability ratio for the historical simulations (1980-2010) is shown with a black error bar, and a reference line at $R = 1$ is shown in all panels. From Longmate et al, (In Revision)

The quantity of interest is then the ratio of the between-model variability to the within-model variability, quantified in terms of the standard deviations ω and τ , denoted:

$$R = \frac{\omega}{\tau} \quad (2)$$

A finding of $R > 1$ (i.e., the between-model variability is much larger than the within-model variability), indicates that a random sampling of ensemble members produces an unbiased multi-model ensemble. However, a finding of $R < 1$ indicates that the variability within the various models is larger than the differences between models, and the choice of the number of ensemble members of a given model for downscaling will impact multi-model assessments. Using the assumptions specified by Equations 1 and 2, standard statistical software can be used to yield maximum-likelihood estimates of the ratio of variances; the Delta method can be used to quantify uncertainty in the ratio in Equation 2.

This investigation constructed CONUS-wide and regional variability ratios R for moving window 30-year averages across the SSP projections from 2015-2100. Figure 20 shows best estimates of R along with a 95% confidence interval, plotted as a function of the average global

mean temperature (GMT) anomaly during each 30-year period for each SSP (where the anomalies are calculated relative to the pre-industrial period). For reference, the variability ratio for historical simulations and its 95% confidence interval is shown at the minimum mean GMT.

All estimates of R (CONUS and regional) have been found to be greater than unity, regardless of GMT value. This finding is unambiguous: downscaling efforts should focus on sampling a broad range of models and efforts to downscale multiple IC ensemble-members per model will be redundant.

Observational Product Analysis

This project is focused on the statistical distributions of temperature and precipitation, with a particular interest in understanding the tails of these distributions. These tails describe extreme events which lead to highly disproportionate damage to operations, infrastructure, and ecosystems (ASCE, 2015). Extreme events are often categorized by return intervals, which are closely related to the probability that a given event will occur during a fixed period of time. A 5% probability that a particular weather phenomenon would occur in any 24-hour period during a given year would be referred to as a 20-year (inverse of the probability of occurrence) daily return event.

The nature of extreme events, including the atmospheric, surface, and sub-surface processes that occur, and the antecedent and subsequent conditions surrounding extreme events, are, by their very definition, rare. Therefore, physical or statistical modeling of such events is most robust when performed in the context of actual events. This is why long observational records that capture extreme events are so valuable and analyses of long return interval events requires data records that exceed the return interval length. For the downscaling of climate model simulations, long data records are especially critical: observations show the actual relationships between the large-scale atmospheric circulations, which steer atmospheric phenomena, and resulting weather systems that produce local effects. Sufficient data are needed to understand the particulars of how extreme events occur to inform the particulars of what climate models need to simulate to be able to predict changes in such events in the future.

Unfortunately, the long-term observational data products that are used to evaluate statistical and dynamical downscaling are, themselves, potentially problematic. The reasons for the issues in observational products can be myriad, but often are the result of raw observational data processing, which can lead to the inadvertent introduction of errors into those products. For example, Figure 21 shows that the time-of-day at which observations have been collected varies throughout the historical record. For precipitation records, this change in time-of-day would not seem to be impactful, since there are no clear diurnal patterns to precipitation that would impact extreme precipitation.

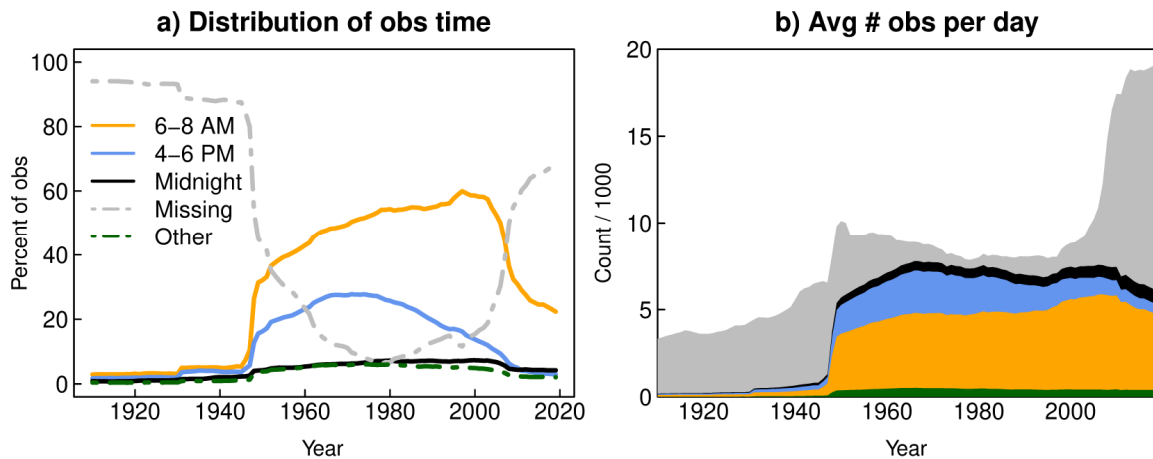


Figure 21: (a) Distribution of precipitation gauge observation times (LST). (b) Average number of observations per day, calculated for each year. Observation time is indicated using the same colors as in (a). From Pierce et al, (2021).

However, previous, widely-utilized hydrometeorological datasets (e.g., Livneh et al, 2013; Livneh et al, 2015) sought to standardize these data by the day it was collected starting at midnight and extending 24 hours forward. The choice to standardize the data led to a need to decide how to attribute precipitation measured at some time other than midnight into the day it was collected or the previous day. A decision was made to assume that precipitation fell uniformly over the 24-hour period when data was collected.

This investigation found that that assumption was problematic for observational dataset construction and model evaluation (Pierce et al, 2021). For example, if a measurement was reported at 6 am, then $\frac{1}{4}$ of the precipitation from the previous day would be combined with $\frac{3}{4}$ of the precipitation of the day in question. In effect, this standardization involved temporal averaging, which necessarily mutes extremes.

This investigation went on to demonstrate how the process of constructing observational data sets for training the empirical-statistical downscaling methods has a direct influence on the downscaling solutions, particularly for the problem of long-period return values of daily precipitation extremes (Risser et al, 2021). As such, when observational products contain biases due to gridding, there are direct impacts to the downscaling solution (e.g., Kannan et al, 2014). The findings of Risser et al, (2021) indicate that the training dataset used for LOCA systematically underestimated extreme precipitation, so LOCA also systematically underestimated extreme precipitation across CONUS. Further systematic analyses are needed to isolate and quantify the impact of the training dataset alone on LOCA solutions.

For the purposes of DoD infrastructure and operations planning, there is a key distinction that must be made pertaining to the interpretation of gridded precipitation data sets. Gridded precipitation can either be interpreted as an areal average over a grid box or as a point estimate (Chen and Knutson, 2008). As described in Gervais et al. (2014), precipitation arises from small-scale structures that are discontinuous in nature, which can lead to large differences between the areal average and point estimate interpretations. Such interpretational differences can

furthermore lead to significant differences in the resulting statistics of precipitation, particularly for extremes (see, e.g., [Gervais et al. 2014](#); [Wehner et al. 2021](#)). GCM-simulated precipitation (and, more generally, precipitation from any dynamical model) is best understood as an areal average regardless of how small the grid boxes are since a physical model can accumulate fluxes through the boundaries of its grid boxes. Measurements of precipitation from weather stations, on the other hand, follow the point estimate interpretation since they represent precipitation collected at an (essentially) infinitesimally small geospatial location. As is demonstrated in [Gervais et al. \(2014\)](#), any gridding method based on interpolation of daily weather station measurements (e.g., bilinear interpolation or inverse distance weighting) will yield a daily gridded precipitation product that should be interpreted as a point estimate.

For these reasons, [Risser et al. \(2019\)](#) produced a data product (henceforth R19) that preserves the extreme statistics of weather station measurements by first calculating extreme statistics at each station and then interpolating these statistics using a data-driven technique to a $\frac{1}{4}^\circ$ (~ 25 km) grid. Furthermore, R19 contains uncertainty quantification in the form of a bootstrap or resampling-based ensemble of estimates. Note that as an interpolation-based method, the R19 product should be interpreted as a point measurement ([Gervais et al. 2014](#)). The primary benefit of the R19 data set is that the spatial aspect of the statistical analysis borrows strength from neighboring stations when estimating the extreme statistics at each station. Specifically, R19 uses second-order nonstationary spatial Gaussian process (GP) models that account for spatial heterogeneities in the climatology of extreme precipitation and enable statistical inference at locations where observations of daily precipitation do not exist. GPs are an extremely popular tool in statistical modeling, because modeling point-referenced, continuously-indexed spatial processes like the climatological coefficients of extreme precipitation because all finite-dimensional distributions are known to be Gaussian, and because the GP is completely specified by a characterization of its first- and second-order properties. They are broadly applied in spatial and environmental statistics as well as machine learning and emulation of complex mathematical models. GPs are a form of “nonparametric” or nonlinear regression, as they characterize a nonlinear relationship between a set of inputs and an output: in our application, the inputs are geographic coordinates and the output (or response) is one of the estimated climatological coefficients of extreme value distribution parameters. Furthermore, with this data-driven approach to estimating spatial coherence and variability of extreme precipitation, R19 automatically takes into account the dominant drivers of the spatial structure of extreme precipitation including orography and major synoptic events (such as hurricanes). It is therefore more physically meaningful than heuristic approaches like bilinear interpolation or inverse-distance weighting, or even “ordinary kriging” using an isotropic Gaussian process, which uses a constant (and circular) spatial length-scale. An additional benefit to the use of the R19 approach is that it uses multiple nearby observations of extreme precipitation to estimate the parameters of an extreme value distribution at a given point. It thereby reduces the uncertainty in such an estimate by incorporating more data into that estimate.

In contrast to R19, the L15 daily gridded product provides estimates of daily precipitation at $1/16^\circ$ spatial resolution over much of North America for 1950–2013. L15 is a widely used data product, and is relevant here because it is used as training data for the LOCA and LOCA2 downscaling method. The L15 data product uses *in situ* measurements of daily total precipitation (over CONUS, the input data are from the Global Historical Climatology Network; [Menne et al.](#)

2012) to create a daily gridded product in two steps. First, for each day, the station measurements are interpolated to a $1/16^\circ$ (~ 6 km) high-resolution grid using the SYMAP algorithm (Shepard 1968, 1984), which is an inverse distance weighting approach. Second, the interpolated data are multiplied by a monthly scaling factor determined by the ratio of its mean monthly baseline climatology (1981–2010) and the mean monthly climatology from the same period of the topographically aware PRISM data product (Daly et al. 1994, 2008).

This investigation intercompared 20-year precipitation return values based on the LOCA dataset, which was trained with L15, and the R19 data product at 163 installations that are affiliated with the Air Force, Army, Coast Guard, Marine Corps, National Guard, or the Navy in the CONUS. As shown in Figure 22, the precipitation return values from LOCA/L15 are significantly lower than those of R19.

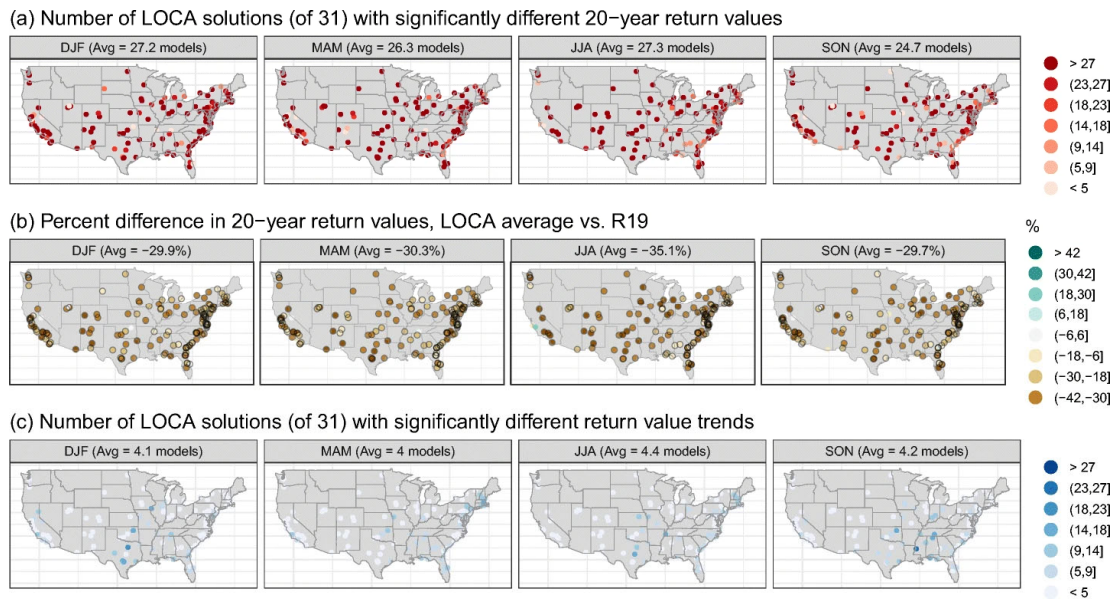


Figure 22: The number of LOCA solutions with significantly different 20-year return value estimates versus R19 (panel a; the average number of models is shown in the strip text), as well as the percent difference between the LOCA-average 20-year return value and the corresponding quantity from R19 (panel b). Locations with a significant difference between R19 and the LOCA average are plotted with a black circle. Panel c shows the number of LOCA solutions for which the 65-year (1951–2005) linear trend in 20-year return value differs from the R19 trend (with the seasonal average number of solutions shown in the strip text). From Risser et al, (2021).

A detailed analysis of the relative importance of time-of-day adjustments and bias due to gridding was also performed in Risser et al (2021). Table (4) shows the fraction of total bias in 20-year precipitation return values across the CONUS due to time-of-day adjustments and bias from gridding. Note that there is some level of convolution to these biases (the sum of the fractional contributions is greater than 1), but the relative contributions are nevertheless instructive.

Season	Total bias (L15)	Frac. of bias TOD	Frac. bias gridding
DJF	-29.4%	0.74	0.37
MAM	-30.9%	0.63	0.43
JJA	-35.7%	0.53	0.51
SON	-30.4%	0.63	0.41

Table 4: Total percent bias in 20-year return values for L15 relative to R19 (second column from the left) averaged over all 0.25° CONUS grid cells with an elevation less than 1500 m. From Risser et al, (2021).

Fortunately, there is a straightforward fix both to the time-of-day biases and gridding biases. For downscaling training, precipitation observations can be analyzed irrespective of the time-of-day that they were collected. For gridding biases, the R19 dataset can be used to scale precipitation return values. As shown in Figure 23, there is a scaling factor for a given precipitation return value to correct for location-specific effects. These vary by location, and Figure 23 displays two scaling factors at two different DoD installations, demonstrating that the scaling factor vs return period is easily parameterized.

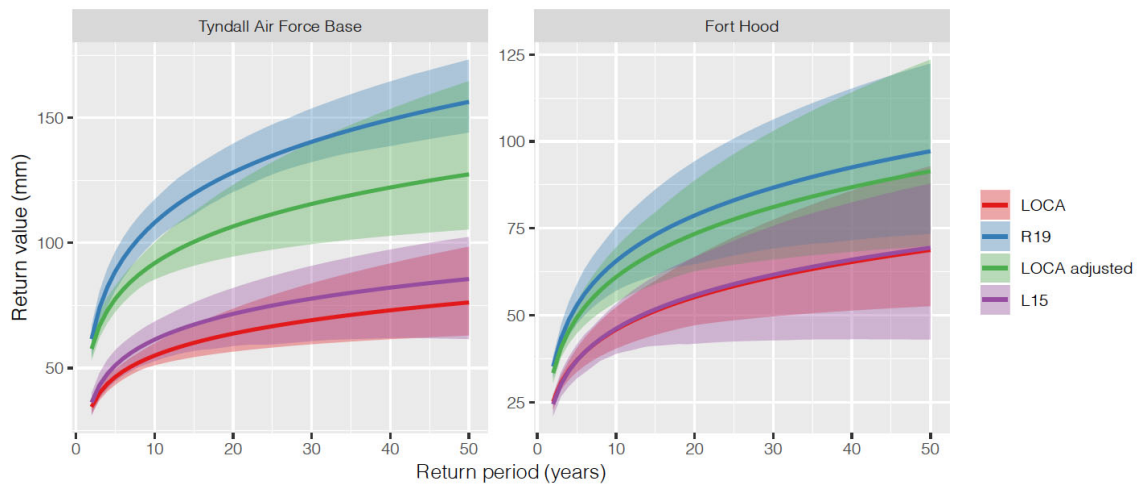


Figure 23: Return values for R19 and the LOCA solutions, before applying the location-specific scaling factor to LOCA at the Tyndall Air Force Base (scaling factor = 1.67) and Fort Hood (scaling factor = 1.33) grid cells. The L15 return values are also shown for reference (recall that the scaling factors are calculated as R19/L15). For R19 and L15, the 95% basic bootstrap confidence band is shown; for the LOCA and LOCA-adjusted, the band represents the multi-model spread (also, the plotted line is the multi-model mean). From Risser et al, (2021).

A final consideration for the use of historical observations, especially for evaluating historical model performance pertains to data homogenization. Due to changes in instrumentation, land-use, and instrumental location/coverage, there are strong and apparent needs to make adjustments to raw station observations, since data artifacts can impact trends.

This investigation has analyzed a wide range of temperature datasets over California and found that homogenized observations produce a very different picture of California warming in the

historical record as compared to nonhomogenized data. Homogenized data include the Berkeley Earth Surface Temperature (BEST) product (Rohde et al, 2020) and nClimDiv (Vose et al, 2014). Nonhomogenized data include L15, ECMWF Reanalysis 20th Century (ERA-20C) (Poli et al, 2016) and NOAA 20th Century Reanalysis Version 3 (Slivinski et al, 2021).

Figure 24 shows the trend in extreme temperature and average high temperature from 1970 to 2005 is for warming in coastal areas with homogenized data, but cooling for nonhomogenized, which this figure shows on its right-hand-side: The red and dark red dots are consistently higher than the orange, yellow, and green dots, showing that non-homogenized historical trends are biased relative to homogenized data across California. This surprising finding highlights the need for a significant amount of care in the use of a model's historical performance as a basis for model weighting based. In point of fact, this finding shows that a model's historical performance against a non-homogenized observational dataset would be very different from that same model's performance against a homogenized observational dataset. The homogenized dataset provides a better measure of model performance by explicitly addressing observational artifacts, but the correction of those artifacts is imperfect and necessarily introduces uncertainty in data products (Weatherhead et al, 1998; Kennedy et al, 2011; Lindau and Venema, 2015).

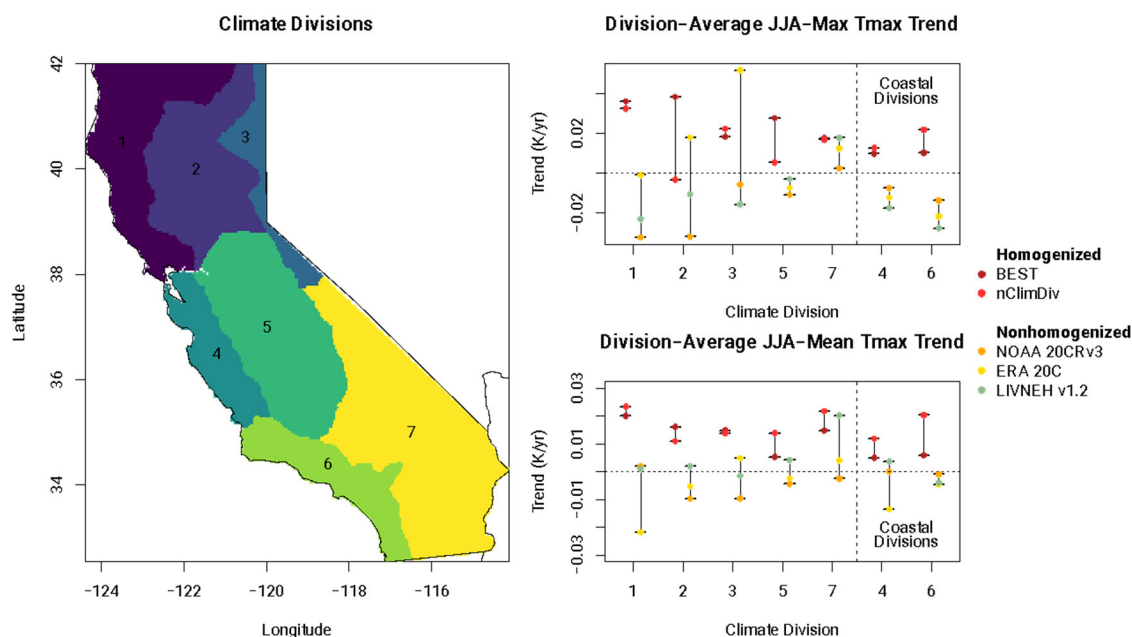


Figure 24: 1970-2005 trends in California climate divisions shown in left from homogenized and nonhomogenized data records. From Charn et al, (2022).

This finding points to the care that is needed for model metric development even over the historical record: spurious trends can arise which homogenization can address, but they have major impacts.

Model Weighting

The 4th National Climate Assessment (NCA4) ([USGCRP, 2017](#)) adopted a model-weighting approach to produce, from the CMIP5 ensemble, regional-scale estimates of change in average temperature and precipitation. That weighting was based on how well each model agrees with observations over the historical record, as well as how independent each model is from other models. From the weighted model results, a consistent story emerged: there was high confidence that the predictions of more warming with the Representative Concentration Pathway 8.5 (RCP8.5) scenario than with the RCP4.5 scenario were robust across the CONUS. At the same time, the weighted models show, albeit with less confidence, more variability in regional precipitation for both RCP8.5 and RCP4.5 than in the historical record.

The approach for model weighting for NCA5 followed a similar conceptual path as NCA4 and Project PI Feldman and Unfunded Co-Performer Arnold regularly contributed to the ad-hoc group that led model weighting for NCA5. That group was organized by Dr. Kyo Lee of the NASA Jet Propulsion Laboratory and included, in addition to Feldman and Arnold, Dr. Ken Kunkel of NOAA, Dr. Keith Dixon of NOAA, Dr. Elias Massoud of the University of California-Berkeley, Dr. Mike Wehner of the Lawrence Berkeley National Laboratory, Dr. Adam Terando of the United States Geological Survey, and Dr. Radley Horton of Columbia University.

With CMIP6 containing both updated models from modeling centers that contributed to CMIP5 and models from centers that had yet to contribute to CMIP, the model weighting working group determined that a multi-pronged approach to testing each CMIP GCM would be needed. First, the simulations from each modeling center were treated generally as independent, while IC ensemble members were not, which was consistent with the findings of Longmate et al (In Revision).

Next, some of the metrics focus on ensuring that a model exhibits realistic atmospheric structure at a global scale. Table 5 shows several variables with global domains. These metrics include seasonally-averaged surface temperature and precipitation over land, surface temperature and precipitation anomalies at low latitudes, top-of-atmosphere (TOA) shortwave and longwave radiative fluxes, vertical temperature/humidity profiles, surface pressure, and GMT surface temperature changes from 1850-1900 to 1985-2015.

Third, some of the metrics focus also on ensuring that a model realistically simulates hydroclimatic conditions over the CONUS. The work supported by this project and published in [Charn et al \(2022\)](#) contributed to the metric formulation. The metrics include (1) the seasonal maximum 5-day total precipitation, (2) the diurnal temperature range, and (3) the warmest day and night. These provide granular detail and test the downscaling performance at the local level. For these variables, each CMIP6 model is compared against the observational datasets indicated in Table 5 to determine that model's skill.

Description	Domain	Source	Years	Reference
Seasonal surface temperature	Global land	HadCru or U. Delaware	1950-2015	Harris et al, 2014
Annual surface temperature anomaly	Low-latitude ocean (50N-50S)	NASA GISS	1950-2015	Reynolds et al, 2002
Seasonal mean precipitation	Global land	REGEN with mask	1950-2016	Sorooshian et al, 2000
Seasonal mean precipitation	Low-latitude ocean (50N-50S)	TRMM, GPCP or equivalent	1988-2015	Huffman et al, 1997
Seasonal TOA shortwave/longwave flux	Global	CERES-EBAF	2000-2015	Loeb et al, 2018
Vertical temperature/humidity profile	Global	AIRS	2002-2015	Aumann et al, 2003
Surface pressure	Global	ERA-5	1980-2015	Muñoz Sabater et al, 2019
Seasonal max. 5-daily total precipitation	CONUS+Canada	L15, Hutchinson	1950-2015	Hopkinson et al, 2012
GMT difference 1850-1900 to 1985-2015	Global	HadCRUTv5	1850-1900 & 1985-2015	Harris et al, 2014
Diurnal temperature range	CONUS+Canada	L15, Hutchinson	1950-2013	L15, Hutchinson et al 2009
Warmest night/day	CONUS+Canada	L15, Hutchinson	1950-2013	L15, Hutchinson et al 2009

Table 5: Metrics to be used by NCA5 to weight CMIP6 models for 21st Century CONUS projections

Future Research Directions and Benefits to DoD

This interim report discusses a set of research activities that are in-progress. To date, the project has produced 5 publications, all of which directly contribute to the project's main objective by helping to test its null hypotheses. These include:

1. Charn et al, (2022) found that the accuracy of the observational products themselves must be first established before assessing model skill (H1)
2. Pierce et al, (2021) and Risser et al (2021) showed the updates needed for LOCA2 (H2) and how climate model skill should be assessed (H1).
3. Longmate et al (In Revision) showed how to sample CMIP6 parsimoniously for downscaling purposes (H1).
4. Rahimi et al (2022) showed how the configuration for an unprecedented set of dynamical downscaling activities that have been completed for CMIP6 (H2).

The project will be focused on completing its tasks and objectives during its stated period of performance (12/31/2019 to 12/31/2023). These include:

1. Completing a comprehensive set of [statistical](#) and [dynamical](#) downscaling calculations.
2. Analyzing results from those calculations to test the null hypotheses of the RC19-1391 project
3. Producing a dataset of projections of 21st Century surface temperature and precipitation over the CONUS for end-users that summarizes the information contained in the CMIP6 multi-model ensemble.

By the end of the period of performance, the project will have produced a large set of statistical and dynamical downscaled solutions from CMIP6 and will provide them to the larger community of end-users that are interested in projections of 21st Century temperature and precipitation. Statistical downscaling solutions from LOCA2 for 27 models have been delivered to NCA5 in the first half of 2022 and dynamical solutions already completed have been delivered to Amazon Web Services for public dissemination also in the first half of 2022. By the end of the project, the research team will request that these data be hosted on the Green Data Oasis (GDO).

Beyond the development of downscaled datasets for end-users, the investigation has taken and will continue to take a close look at the downscaling process chain to provide clarity for end-users as to the uncertainty in these datasets and how information from terabytes of simulations can be distilled and adopted for climate impacts analysis. To that end, the project has the following interim findings:

1. The observational data record can contain artifacts arising from assumptions made in converting raw observations to observational products. Even though observational products are needed to test model performance, these assumptions must be addressed formally before developing downscaling solutions. Artifacts in widely-utilized precipitation and temperature data products have been uncovered, and they yield erroneous downscaling solutions if they are used incorrectly in assessing CMIP6 model skill over the historical record.
2. The divergence between statistical and/or dynamical downscaling solutions in future projections shows where there are processes that can produce non-stationary effects that can impact surface temperature and precipitation projections. The divergence in solutions does not point necessarily to a requirement to use one form of downscaling or another.
3. While the CMIP6 archive contains dozens of models, hundreds of experiments, and thousands of ensemble member integrations, the downscaling of the CMIP6 does not have to be exhaustive. Limited calculations, developed by sampling model ensemble members, is generally sufficient for developing a suite of downscaled projections that captures the range of the ensemble.

4. Location-specific analyses of different climate and weather phenomena are needed to produce insights into the atmospheric and land-surface processes that would contribute to the potential for nonstationarity in downscaling solutions.
5. Bias-correction of models is appropriate for ensuring products are relevant to end-users, but if the bias correction for a model is larger than the centennial-length projected change, it may not be prudent to expend computational and personnel resources to downscale that model.
6. To date, the investigation has used conventional statistical and dynamical downscaling methodologies. There may be a role for artificial intelligence (AI), especially to accelerate the production of dynamical downscaling solutions through whole model or individual parameterization emulation of WRF. The project is working with topical experts in the AI research community to determine the risks and scientific rewards for undertaking these novel approaches for emulation.

A key deliverable from this work will be new projections of future climate-changed conditions of temperature and precipitation at finer spatial scales than are most often available for climate vulnerability assessments. The project will use the aforementioned findings to develop a condensed and easy-to-use set of downscaled temperature and precipitation fields over the CONUS and their uncertainties to end-users, wherein uncertainties from nonstationary effects are explicitly estimated.

The data, especially the condensed data, can be widely utilized by the scientific and DoD user communities. Additionally, one specific use-case for this deliverable is that it will help characterize risk in specific watersheds around potentially vulnerable DoD installations. Additionally, the LOCA2-derived future climatology from this project may help update the findings in the FY18 “National Defense Appropriations Act (NDAA; Public Law 115-91) Section 335 Evaluation of Climate Change Effects on Army Installations” [White, K.D. et al., USACE Washington, DC 2018] delivered to Congress in January 2019 in the “Report on Effects of a Changing Climate to the Defense Department,” from the Office of the Under Secretary of Defense for Acquisition and Sustainment, to specifically incorporate CMIP6 model results.

In the latter half of the investigation, once the project has developed and downscaling solutions over the CONUS for CMIP6, the project will be engaging with DoD end-users, including with other SERDP/ESTCP projects, and NCA5 to perform case-studies and develop insights into best practices for the provisioning of these simulations to end-users.

Literature Cited

ASCE [American Society of Civil Engineers] (2015). *Adapting Infrastructure and Civil Engineering Practice to a Changing Climate*. Committee on Adaptation to a Changing Climate, Olsen, J. Rolf, ed. Technical Report, American Society of Civil Engineers, Reston, VA. doi: [10.1061/9780784479193](https://doi.org/10.1061/9780784479193).

Aumann, H.H. et al., (2003) AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems, *IEEE Transactions on Geoscience and Remote Sensing*, **41**(2), 253-264, doi: [10.1109/TGRS.2002.808356](https://doi.org/10.1109/TGRS.2002.808356).

Bengtsson, L., K.I. Hodges, and E. Roeckner (2006) Storm Tracks and Climate Change. *Journal of Climate*, **19**, 3518–3543, doi:[10.1175/JCLI3815.1](https://doi.org/10.1175/JCLI3815.1).

Bruyère, C.L., Done, J.M., Holland, G.J. et al. (2014), Bias corrections of global models for regional climate simulations of high-impact weather. *Climate Dynamics*, **43**, 1847–1856, doi: [10.1007/s00382-013-2011-6](https://doi.org/10.1007/s00382-013-2011-6).

Charn, A. B., O'Brien, T. A., Risser, M. D., Longmate, J. M., & Feldman, D. R. (2022). Sign of observed California temperature trends depends on data set homogenization: Implications for weighting and downscaling. *Geophysical Research Letters*, **49**, e2022GL099186, doi:[10.1029/2022GL099186](https://doi.org/10.1029/2022GL099186).

Charney, J. et al, National Research Council. (1979). *Carbon Dioxide and Climate: A Scientific Assessment*. Washington, DC: The National Academies Press. doi:[10.17226/12181](https://doi.org/10.17226/12181).

Chen, C., & Knutson, T. (2008). On the Verification and Comparison of Extreme Rainfall Indices from Climate Models, *Journal of Climate*, **21**(7), 1605-1621, doi: [10.1175/2007JCLI1494.1](https://doi.org/10.1175/2007JCLI1494.1).

Cinquini, L. et al (2014) The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data, *Future Generation Computer Systems*, **36**, 400-417, doi: [10.1016/j.future.2013.07.002](https://doi.org/10.1016/j.future.2013.07.002).

Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**(2):140–158, doi: [10.1175/1520-0450\(1994\)033<0140:ASTMFM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2).

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the Conterminous United States. *International Journal of Climatology*, **28**(15):2031–2064, doi:[10.1002/joc.1688](https://doi.org/10.1002/joc.1688).

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert. (2012) HESS opinions: ‘Should we apply bias correction to global and regional climate model data?’ *Hydrological and Earth System Sciences*, **16**: 3391–3404. doi: [10.5194/hess-16-3391-2012](https://doi.org/10.5194/hess-16-3391-2012).

Eyring, V., et al., (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Devevelopment*, **9**, 1937-1958, doi:[10.5194/gmd-9-1937-2016](https://doi.org/10.5194/gmd-9-1937-2016).

Feldman, D.R., J. Tadić, A. Schwarz, W. Arnold (2021) Establishing a Range of Downscaled Precipitation Estimates in California for Planning in the Face of Climate Change, *Journal of Water Resources Planning and Management*, **147**(9), doi: [10.1061/\(ASCE\)WR.1943-5452.0001410](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001410).

Forster, P.M., (2016) Inference of Climate Sensitivity from Analysis of Earth's Energy Budget, *Annual Review of Earth and Planetary Sciences*, **44**, 85-106, doi:[10.1146/annurev-earth-060614-105156](https://doi.org/10.1146/annurev-earth-060614-105156).

Gervais M, Tremblay LB, Gyakum JR, Atallah E (2014) Representing extremes in a daily gridded precipitation analysis over the United States: Impacts of station density, resolution, and gridding methods. *Journal of Climate*, **27**(14):5201–5218, doi: [10.1175/JCLI-D-13-00319.1](https://doi.org/10.1175/JCLI-D-13-00319.1).

Giorgi, F., C. Jones, and G. R. Asrar, (2009) Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bull.*, **58**, 175–183.

Haarsma, R.J., et al., (2016) High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, *Geoscientific Model Devevelopment*, **9**, 4185-4208, doi:[10.5194/gmd-9-4185-2016](https://doi.org/10.5194/gmd-9-4185-2016).

Harris, I., Jones, P., Osborn, T. and Lister, D. (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, **34**: 623-642. doi:[10.1002/joc.3711](https://doi.org/10.1002/joc.3711).

Hawkins, E., & Sutton, R. (2009). The Potential to Narrow Uncertainty in Regional Climate Predictions, *Bulletin of the American Meteorological Society*, **90**(8), 1095-1108, doi:[10.1175/2009BAMS2607.1](https://doi.org/10.1175/2009BAMS2607.1).

Hidalgo, H., M. D. Dettinger, and D. R. Cayan, (2008) Downscaling with constructed analogues: daily precipitation and temperature fields over the United States. California Energy Commission report CEC-500-2007-123. 48 pp.

Hopkinson, R. F., Hutchinson, M. F., McKenney, D. W., Milewska, E. J., & Papadopol, P. (2012). Optimizing Input Data for Gridding Climate Normals for Canada, *Journal of Applied Meteorology and Climatology*, **51**(8), 1508-1518, doi: [10.1175/JAMC-D-12-018.1](https://doi.org/10.1175/JAMC-D-12-018.1).

Hu, Y., L. Tao, J. Liu (2013) Poleward expansion of the hadley circulation in CMIP5 simulations, *Advances in Atmospheric Sciences*, **30**, 790-795, doi:[10.1007/s00376-012-2187-4](https://doi.org/10.1007/s00376-012-2187-4).

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., & Schneider, U. (1997). The Global Precipitation Climatology Project (GPCP)

Combined Precipitation Dataset, *Bulletin of the American Meteorological Society*, **78**(1), 5-20, doi: [10.1175/1520-0477\(1997\)078<0005:TGPCPG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2).

Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E., & Papadopol, P. (2009). Development and Testing of Canada-Wide Interpolated Spatial Models of Daily Minimum–Maximum Temperature and Precipitation for 1961–2003, *Journal of Applied Meteorology and Climatology*, **48**(4), 725-741, doi: [10.1175/2008JAMC1979.1](https://doi.org/10.1175/2008JAMC1979.1).

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Jeevanjee, N. & D.M. Romps (2018) Mean precipitation change from a deepening troposphere, *Proceedings of the National Academy of Sciences of the USA*, **115**, 11465-11470, doi:[10.1073/pnas.1720683115](https://doi.org/10.1073/pnas.1720683115).

Jones, C., F. Giorgi, and G. Asrar, (2011) The Coordinated Regional Downscaling Experiment: CORDEX—An international downscaling link to CMIP5. CLIVAR Exchanges, Vol. 56, No. 16, International CLIVAR Project Office, Southampton, United Kingdom, 34–40.

Kain, J. S. (2004). The Kain–Fritsch Convective Parameterization: An Update, *Journal of Applied Meteorology*, **43**(1), 170-181, doi: [10.1175/1520-0450\(2004\)043<0170:TKCPAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2).

Kannan, S., Ghosh, S., Mishra, V., and Salvi, K. (2014), Uncertainty resulting from multiple data usage in statistical downscaling, *Geophysical Research Letters*, **41**, 4013– 4019, doi:[10.1002/2014GL060089](https://doi.org/10.1002/2014GL060089).

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., & Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability, *Bulletin of the American Meteorological Society*, **96**(8), 1333-1349, doi: [10.1175/BAMS-D-13-00255.1](https://doi.org/10.1175/BAMS-D-13-00255.1).

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., and Saunby, M. (2011), Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization, *Journal of Geophysical Research*, **116**, D14104, doi:[10.1029/2010JD015220](https://doi.org/10.1029/2010JD015220).

Knutti, R., G. Hegerl (2008), The equilibrium sensitivity of the Earth's temperature to radiation changes, *Nature Geoscience*, **1**, 735-743, doi:[10.1038/ngeo337](https://doi.org/10.1038/ngeo337).

- Knutti, R., M. A. A. Rugenstein, G.C. Hegerl, (2017), Beyond equilibrium climate sensitivity, *Nature Geoscience*, **10**, 727-736, doi:[10.1038/ngeo3017](https://doi.org/10.1038/ngeo3017).
- Kunkel, K.E., D.R. Easterling, (2016) Climate Scenarios for the Fourth National Climate Assessment and the Sustained Assessment Process, American Geophysical Union, Fall Meeting 2016, abstract #GC31G-1180.
- Lavin-Gullon, A, Fernandez, J, Bastin, S, *et al.* (2021) Internal variability versus multi-physics uncertainty in a regional climate model. *International Journal of Climatology*, 41 (Suppl. 1): E656– E671. doi:[10.1002/joc.6717](https://doi.org/10.1002/joc.6717).
- Lehner, F. et al, (2020) Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6, *Earth System Dynamics*, **11**, 491–508, 2020, doi: [10.5194/esd-11-491-2020](https://doi.org/10.5194/esd-11-491-2020).
- Lindau, R. and Venema, V. (2016), The uncertainty of break positions detected by homogenization algorithms in climate records. *International Journal of Climatology*, 36: 576-589. <https://doi.org/10.1002/joc.4366>
- Livneh, B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer, and D.P. Lettenmaier (2013) A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *Journal of Climate*, **26**, 9384–9392, doi:[10.1175/JCLI-D-12-00508.1](https://doi.org/10.1175/JCLI-D-12-00508.1).
- Livneh, 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 US, and Southern Canada 1950–2013. *Scientific data*, **2**(1):1–12, doi: [10.1038/sdata.2015.42](https://doi.org/10.1038/sdata.2015.42).
- Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., & Kato, S. (2018). Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, *Journal of Climate*, **31**(2), 895-918, doi: [10.1175/JCLI-D-17-0208.1](https://doi.org/10.1175/JCLI-D-17-0208.1).
- Longmate, J.M., M.D. Risser, D.R. Feldman (2022) Prioritizing the Selection of CMIP6 Model Ensemble Members for Downscaling Projections of CONUS Temperature and Precipitation, *Climate Dynamics*, (In Revision)
- Lu, J., G. A. Vecchi, and T. Reichler (2007), Expansion of the Hadley cell under global warming, *Geophysical Research Letters*, 34, L06805, doi:[10.1029/2006GL028443](https://doi.org/10.1029/2006GL028443).
- Maher, N., S. Milinski, R. Ludwig (2021) Large ensemble climate model simulations: introduction, overview, and future prospects for utilising multiple types of large ensemble, *Earth System Dynamics*, **12**, 401-418, doi: [10.5194/esd-12-401-2021](https://doi.org/10.5194/esd-12-401-2021).

- Mankin, J. S., Lehner, F., Coats, S., & McKinnon, K. A. (2020). The value of initial condition large ensembles to robust adaptation decision-making. *Earth's Future*, **8**, e2012EF001610. doi:[10.1029/2020EF001610](https://doi.org/10.1029/2020EF001610).
- Mauritsen, T., et al. (2012). Tuning the climate of a global model, *Journal of Advances in Modeling Earth Systems*, **4**, doi:[10.1029/2012MS000154](https://doi.org/10.1029/2012MS000154).
- Mearns, L. O., et al (2012) The North American Regional Climate Change Assessment Program: Overview of phase I results. *Bulletin of the American Meteorological Society*, **93**, 1337–1362, doi:[10.1175/BAMS-D-11-00223.1](https://doi.org/10.1175/BAMS-D-11-00223.1).
- Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the Global Historical Climatology Network-Daily database. *Journal of Atmospheric and Oceanic Technology*, **29**(7):897–910, doi: [10.1175/JTECH-D-11-00103.1](https://doi.org/10.1175/JTECH-D-11-00103.1).
- Muñoz Sabater, J., (2019): ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 23-Feb-2020), [10.24381/cds.e2161bac](https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview).
- NRC: National Research Council (U.S.). Geophysics Study Committee. Scientific Basis of Water-resource Management. (National Academy Press, 1982).
- O'Neill, B.C., Kriegler, E., Riahi, K. et al. (2014) A new scenario framework for climate change research: the concept of shared socioeconomic pathways, *Climatic Change*, **122**, 387, doi:[10.1007/s10584-013-0905-2](https://doi.org/10.1007/s10584-013-0905-2).
- O'Neill, B.C., et al. (2016) The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geoscientific Model Development*, **9**, 3461–3482, doi:[10.5194/gmd-9-3461-2016](https://doi.org/10.5194/gmd-9-3461-2016).
- Office of the Under Secretary of Defense for Acquisition and Sustainment (2019), Report on Effects of a Changing Climate to the Department of Defense, <https://media.defense.gov/2019/Jan/29/2002084200/-1/-1/1/CLIMATE-CHANGE-REPORT-2019.PDF>
- Pendergrass, A.G. and D.L. Hartmann (2014) Changes in the Distribution of Rain Frequency and Intensity in Response to Global Warming. *Journal of Climate*, **27**, 8372–8383, doi:[10.1175/JCLI-D-14-00183.1](https://doi.org/10.1175/JCLI-D-14-00183.1).
- Pierce, D. W., D. R. Cayan, and B. L. Thrasher, (2014) Statistical downscaling using Localized Constructed Analogs (LOCA). *Journal of Hydrometeorology*, **15**, 2558–2585, doi:[10.1175/JHM-D-14-0082.1](https://doi.org/10.1175/JHM-D-14-0082.1).
- Pierce, D. W., D. R. Cayan, J. F. Kalansky. (Scripps Institution of Oceanography). 2018. Climate, Drought, and Sea Level Rise Scenarios for the Fourth California Climate Assessment. California's Fourth Climate Change Assessment, California Energy Commission. Publication number: CCCA4-CEC-2018-006.

Pierce, D. W., Su, L., Cayan, D. R., Risser, M. D., Livneh, B., & Lettenmaier, D. P. (2021). An Extreme-Preserving Long-Term Gridded Daily Precipitation Dataset for the Conterminous United States, *Journal of Hydrometeorology*, 22(7), 1883-1895, doi: [10.1175/JHM-D-20-0212.1](https://doi.org/10.1175/JHM-D-20-0212.1).

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., & Fisher, M. (2016). ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, *Journal of Climate*, 29(11), 4083-4097, doi: [10.1175/JCLI-D-15-0556.1](https://doi.org/10.1175/JCLI-D-15-0556.1).

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., Peckham, S. E., Grell, G. A., Michalakes, J., Trahan, S., Benjamin, S. G., Alexander, C. R., Dimego, G. J., Wang, W., Schwartz, C. S., Romine, G. S., Liu, Z., Snyder, C., Chen, F., Barlage, M. J., Yu, W., & Duda, M. G. (2017). The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions, *Bulletin of the American Meteorological Society*, 98(8), 1717-1737, doi:[10.1175/BAMS-D-15-00308.1](https://doi.org/10.1175/BAMS-D-15-00308.1).

Rahimi, S., W. Krantz, Y.-H. Lin, B. Bass, N. Goldenson, A. Hall, Z.J. Lebo, and J. Norris (2022). Evaluation of a reanalysis-driven configuration of WRF4 over the western United States from 1980 to 2020. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035699, doi:[10.1029/2021JD035699](https://doi.org/10.1029/2021JD035699).

Raymond, C., Horton, R.M., Zscheischler, J. et al. (2020), Understanding and managing connected extreme events. *Nature Climate Change*, 10, 611–621. doi: 10.1038/s41558-020-0790-4.

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An Improved In Situ and Satellite SST Analysis for Climate, *Journal of Climate*, 15(13), 1609-1625, doi: [10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2).

Risser, M.D., Paciorek, C.J., Wehner, M.F. et al. A probabilistic gridded product for daily precipitation extremes over the United States. *Climate Dynamics*, 53, 2517–2538 (2019). doi:10.1007/s00382-019-04636-0.

Risser, M.D., D.R. Feldman, M.F. Wehner, D.W. Pierce, D.R. Cayan, J.R. Arnold, (2021), Identifying and correcting biases in localized downscaling estimates of daily precipitation return values, *Climatic Change*, 169, 33, doi:[10.1007/s10584-021-03265-z](https://doi.org/10.1007/s10584-021-03265-z).

Rohde, R. A. and Hausfather, Z. (2020) The Berkeley Earth Land/Ocean Temperature Record, *Earth System. Science Data*, 12, 3469–3479, doi: [10.5194/essd-12-3469-2020](https://doi.org/10.5194/essd-12-3469-2020).

Ruane A C, Teichmann C, Arnell N W, Carter T R, Ebi K L, Frieler K, Goodess C M, Hewitson B, et al. (2016). The Vulnerability, Impacts, Adaptation and Climate Services Advisory Board (VIACS AB v1.0) contribution to CMIP6. *Geoscientific Model Development* 9 (9): 3493-3515. DOI:[10.5194/gmd-9-3493-2016](https://doi.org/10.5194/gmd-9-3493-2016).

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, *Journal of Geophysical Research: Atmospheres*, **118**, 10,884–10,906, doi:[10.1002/jgrd.50843](https://doi.org/10.1002/jgrd.50843).

Sanderson, B. M., Knutti, R., and Caldwell, P. (2015) A representative democracy to reduce interdependency in a multimodel ensemble, *Journal of Climate*, **28**, 5171–5194, doi:[10.1175/JCLI-D-14-00362.1](https://doi.org/10.1175/JCLI-D-14-00362.1).

Sanderson, B.M., Wehner, M., Knutti, R. (2017). Skill and independence weighting for multi-model assessments, *Geoscientific Model Development*, **10**, 2379-2395, doi:[10.5194/gmd-10-2379-2017](https://doi.org/10.5194/gmd-10-2379-2017).

Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N. L., Andrejczuk, M., et al. (2020). Implementation of U.K. Earth system models for CMIP6. *Journal of Advances in Modeling Earth Systems*, **12**, e2019MS001946. doi: [10.1029/2019MS001946](https://doi.org/10.1029/2019MS001946).

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: *Proceedings of the 1968 23rd ACM national conference*. pp 517–524.

Shepard DS (1984) Computer mapping: The SYMAP interpolation algorithm. In: *Spatial statistics and models*. Springer, pp 133–145

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., & Huang, X.-Y. (2019). A Description of the Advanced Research WRF Model Version 4, NCAR Technical Note No. NCAR/TN-556+STR, doi: 10.5065/1DFH-6P97.

Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., & Wyszynski, P. (2021). An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3, *Journal of Climate*, **34**(4), 1417-1438, doi: [10.1175/JCLI-D-20-0505.1](https://doi.org/10.1175/JCLI-D-20-0505.1).

Sorooshian, S., Hsu, K., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall, *Bulletin of the American Meteorological Society*, **81**(9), 2035-2046, doi: [10.1175/1520-0477\(2000\)081<2035:EOPSSE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2).

Soukissian, T.H., Tsalis, C. (2015), The effect of the generalized extreme value distribution parameter estimation methods in extreme wind speed prediction. *Nature: Hazards* **78**, 1777–1809, doi:[10.1007/s11069-015-1800-0](https://doi.org/10.1007/s11069-015-1800-0).

Szopa, S., V. Naik, B. Adhikary, P. Artaxo, T. Berntsen, W.D. Collins, S. Fuzzi, L. Gallardo, A. Kiendler Scharr, Z. Klimont, H. Liao, N. Unger, and P. Zanis, 2021: *Short-Lived Climate Forcers. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Taylor K.E., (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, **106**(D7): 7183–7192, doi: [10.1029/2000JD900719](https://doi.org/10.1029/2000JD900719).

Taylor, K.E., R.J. Stouffer, and G.A. Meehl, (2012) An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**, 485–498, doi:[10.1175/BAMS-D-11-00094.1](https://doi.org/10.1175/BAMS-D-11-00094.1).

Thorne, James H., Joseph Wraithwall, Guido Franco. (2018). California's Changing Climate 2018. *California's Fourth Climate Change Assessment*, California Natural Resources Agency.

Thornton, Peter E, Thornton, Michele M, Mayer, Benjamin W, Wilhelmi, Nate, Wei, Yaxing, Devarakonda, Ranjeet, and Cook, Robert B. *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2.*. United States: N. p., 2014. Web.

Tiedtke, M. (1989) A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large Scale Models. *Monthly Weather Review*, **117**, 1779-1800. doi:[10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2).

Ulbrich, U., J.G. Pinto, H. Kupfer, G.C. Leckebusch, T. Spanghel, and M. Meyers (2008) Changing Northern Hemisphere Storm Tracks in an Ensemble of IPCC Climate Change Simulations. *Journal of Climate*, **21**, 1669–1679, doi:[10.1175/2007JCLI1992.1](https://doi.org/10.1175/2007JCLI1992.1).

USGCRP, 2017: *Climate Science Special Report: Fourth National Climate Assessment, Volume I* [Wuebbles, D.J., D.W. Fahey, K.A. Hibbard, D.J. Dokken, B.C. Stewart, and T.K. Maycock (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 470 pp, doi: [10.7930/J0J964J6](https://doi.org/10.7930/J0J964J6).

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Lamarque, J.-F., Masui, To., Meinshausen, M., Nakicenovic, N., Smith, S., Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, **109**, 5-3, doi:[10.1007/s10584-011-0148-z](https://doi.org/10.1007/s10584-011-0148-z).

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams, C. N., Jr., Fenimore, C., Gleason, K., & Arndt, D. (2014). Improved Historical Temperature and Precipitation Time Series for U.S. Climate Divisions, *Journal of Applied Meteorology and Climatology*, **53**(5), 1232-1251, doi: [10.1175/JAMC-D-13-0248.1](https://doi.org/10.1175/JAMC-D-13-0248.1).

Walton, D., N. Berg, D. Pierce, E. Maurer, A. Hall, Y.-H. Lin, S. Rahimi, and D. Cayan. (2020) Understanding differences in California climate projections produced by dynamical and statistical downscaling.” *Journal of Geophysical Research: Atmospheres*, **125** (19): e2020JD032812, doi: [10.1029/2020JD032812](https://doi.org/10.1029/2020JD032812).

Weatherhead, E. C., et al. (1998), Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *Journal of Geophysical Research*, **103**(D14), 17149– 17161, doi:[10.1029/98JD00995](https://doi.org/10.1029/98JD00995).

Wehner MF, Lee J, Risser MD, Ullrich P, Gleckler P, Collins WD (2021) Evaluation of extreme sub- daily precipitation in high-resolution global climate model simulations. *Philosophical Transactions of the Royal Society*, **379**. <https://doi.org/10.1098/rsta.2019.0545>.

Wrzesien, M. L., & Pavelsky, T. M. (2020). Projected Changes to Extreme Runoff and Precipitation Events From a Downscaled Simulation Over the Western United States. *Frontiers in Earth Science*, **7**. doi:[10.3389/feart.2019.00355](https://doi.org/10.3389/feart.2019.00355).

Yin, J. H. (2005), A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophysical Research Letters*, **32**, L18701, doi:[10.1029/2005GL023684](https://doi.org/10.1029/2005GL023684).

Zhang, C., Wang, Y., & Hamilton, K. (2011). Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme, *Monthly Weather Review*, **139**(11), 3489-3513, doi: [10.1175/MWR-D-10-05091.1](https://doi.org/10.1175/MWR-D-10-05091.1).