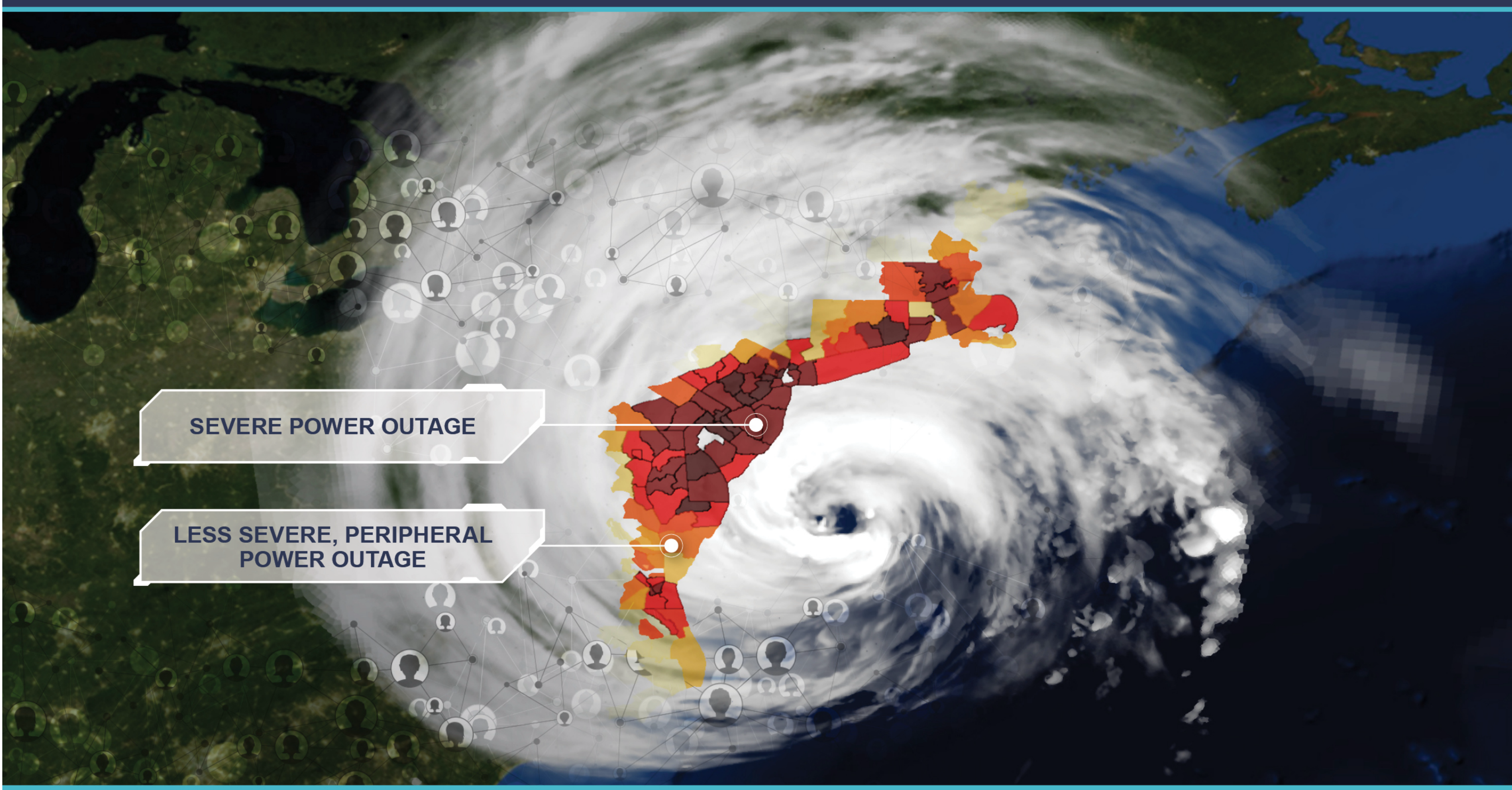
DOE/SC-2021

Understanding Decision-Relevant Regional Climate Data Products

WORKSHOP REPORT



U.S. Department of ENERGY | Office of Science

Biological and Environmental Research Program

October 2024

Understanding Decision-Relevant Regional Climate Data Products Workshop

October 2024

Convened by U.S. Department of Energy
Office of Science, Biological and Environmental Research

P Ullrich, Lawrence Livermore National Laboratory (LLNL)
D Feldman, Lawrence Berkeley National Laboratory (LBNL)
S Abdelrahim, Federal Emergency Management Agency (FEMA)
D Barrie, National Oceanic and Atmospheric Administration (NOAA)
S Basile, U.S. Global Change Research Program (USGCRP)
K Dixon, NOAA
R Joseph, U.S. Department of Energy, Office of Science
D Herring, NOAA
K Hiers, Strategic Environmental Research and Development Program
H Lee, National Aeronautics and Space Administration, Jet Propulsion Laboratory
F Lipshultz, USGCRP
T Spero, U.S. Environmental Protection Agency (EPA)

Scientific Steering Committee

Recommended Citation: U.S. DOE. 2024. *Understanding Decision-Relevant Regional Data Products Workshop Report*, DOE/SC-2021, Department of Energy's Biological and Environmental Research Program, U.S. Department of Energy, Office of Science. https://doi.org/10.2172/2474992

About BER: The BER program advances fundamental research and scientific user facilities to support U.S. Department of Energy missions in scientific discovery and innovation, energy security, and environmental responsibility. BER seeks to understand biological, biogeochemical, and physical principles needed to predict a continuum of processes occurring across scales, from molecular and genomics-controlled mechanisms to environmental and Earth system change. BER advances understanding of how Earth's dynamic, physical, and biogeochemical systems (atmosphere, land, oceans, sea ice, and subsurface) interact and affect future Earth system and environmental change. This research improves Earth system model predictions and provides valuable information for energy and resource planning.

Cover Image: A hypothetical tropical cyclone makes landfall near Philadelphia, as simulated using the U.S. Department of Energy's (DOE) Energy Exascale Earth System Model (E3SM) in its Simple Cloud Resolving E3SM Atmosphere Model (SCREAM) configuration on a regionally refined grid.

- The simulation is by Colin Zarzycki, Pennsylvania State University. Visualization is by Paul Ullrich, Lawrence Livermore National Laboratory, and the University of California Davis. They were supported by the DOE Office of Science, Office of Biological and Environmental Research program under Award Number DE-SC0016605.
- Outage estimates are by Karthik Balaguru and Julian Rice of Pacific Northwest National Laboratory. They were supported by the DOE Office of Science Biological and Environmental Research as part of the collaborative, multiprogram Integrated Coastal Modeling (ICoM) project.
- Work performed by Ullrich is under the auspices of DOE by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.
- Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.
- The research used computational resources from the National Energy Research Scientific Computing Center (NERSC), a DOE User Facility supported by the Office of Science under contract DE-AC02-05CH11231.
- Work performed at Lawrence Berkeley National Laboratory is supported by the U.S. Department of Energy under Contract No. DE-AC02–05CH11231.

Executive Summary

A broad community of climate adaptation practitioners, stakeholders and policymakers rely on historical reconstructions and future projections of local to regional climate. To be of value to these users, climate data must be credible, salient, and authoritative (Cash et al. 2002). Namely, data must be consistent with our physical understanding of the global Earth system, must be relevant for informing the decision-making process, and must be backed by expert judgment. As more and more data products have become available, multiple challenges have emerged around the production, evaluation, selection, and use of these data products. Consequently, to ensure crucial decisions leverage the best possible historical and future physical climate data, there is a pressing need to develop a coordinated national climate data strategy that is inclusive of all relevant communities of practice.

In response to this need, the “Understanding Decision-Relevant Regional Climate Data Projections” workshop was held in-person and virtually from November 14-16, 2023, in Berkeley, California. This workshop was coordinated by the U.S. Global Change Research Program (USGCRP) Interagency Group on Integrative Modeling (IGIM) and the Federal Adaptation and Resilience Group (FARG). Participation came from most major U.S. federal agencies and their partners who are involved in the production and dissemination of regional climate data products, including the U.S. Department of Energy (USDOE), the National Oceanic and Atmospheric Administration (NOAA), the National Aeronautics and Space Administration (NASA), the U.S. Environmental Protection Agency (EPA), Federal Emergency Management Agency (FEMA), U.S. Geological Survey (USGS), U.S. Bureau of Reclamation (USBR), and the Department of Defense’s (DOD) Strategic Environmental Research and Development Program and Environmental Security Technology Certification Program (SERDP and ESTCP). The workshop brought together a wide range of researchers, data producers, end-users, and interagency representatives to understand the current state of the nation’s decision-relevant regional climate projections and carry that understanding forward to enable the development of guidelines for the usage and evaluation of such projections. Numerous approaches for generating regional climate data were discussed, including statistical downscaling, dynamical downscaling, hybrid downscaling, regionally refined global modeling and artificial intelligence. This effort provided a forum for sharing knowledge, establishing common ground, and moving towards the development of a community of practice around decision-relevant data.

The workshop was organized into four sessions focused on 1) data production; 2) data use; 3) data evaluation; and 4) emerging topics. The session on data production featured 10 talks from a variety of data producers, representing multiple federal agencies and academic research groups, followed by breakout sessions that sought to frame the needs of a community of practice. The session on data use featured two panels, each with four panelists presenting brief talks on topics related to how they employ climate data and their perceptions of gaps among existing data products. The session on data evaluation again featured two panels, each with four panelists presenting brief talks related to ascertaining credibility of climate data. The final session on emerging topics featured 11 technically oriented talks on topics related to climate data, including bias correction, model weighting, ensembles, and performance across scales. For each of the first three sessions, there was an accompanying breakout discussion which featured a mix of participants who addressed key questions related to that session and the context of the broader workshop theme.

This executive summary provides a high-level synthesis of discussions at the workshop, focusing on the outstanding challenges identified during the workshop and also potential deliverables from a nascent community of practice to address these challenges.

Challenges for the Decision-Relevant Climate Data Community

Building a common vocabulary: In the course of the workshop, it became clear that a common vocabulary across related communities is needed. Terms such as “extreme event” can have different meanings depending on the needs of particular end-users and the impacts they are considering. The workshop itself was framed around “decision-relevant” or “actionable” data products, but the regional extent, spatial resolution, and temporal resolution for a product to be considered decision-relevant varies depending on the decision being made. “Uncertainty” and “confidence” also emerged as terms that are widely employed in the climate data space, but precise, quantitative definitions of these terms are rarely provided.

Filling data gaps: Despite rapid growth in the number of climate data products, conspicuous gaps remain. For instance, although some statistically downscaled products have global coverage, higher-resolution coverage of areas outside the contiguous United States (OCONUS), including Alaska, Puerto Rico, and island territories, is still needed. Additionally, few high-temporal-resolution (hourly to sub-hourly) data products are available even in the contiguous United States (CONUS) (not to mention OCONUS), despite being needed for many applications (e.g., evaluating sufficiency of storm sewers and projections of renewable energy production). Many opportunities exist for addressing these gaps through new simulations or innovative downscaling methods.

Cataloging and characterizing decision-relevant climate data products: Dozens of regional climate data products have emerged in the past decade at local-to-global scales. They exhibit a variety of spatial and temporal resolutions and feature a variety of climate variables. However, in the absence of a central catalog of data products, end-users and researchers have largely relied on word of mouth and Internet searches to identify relevant data products. Consequently, other, equally relevant products have likely been underused or unused. A catalog of data products, their characteristics, relevant expert guidance, and evaluation metrics could benefit all members of the climate data community, and enable the identification of gaps and synergies among presently available products.

Provisioning common-format, decision-relevant climate data products: Related to the aforementioned challenge of cataloging these products, additional challenges exist in provisioning these data. Three bottlenecks generally stymie data producers interested in provisioning their data to a broader audience: firewalls at the data source, access restrictions and data provisioning support requirements. The sheer size of these data products creates provisioning challenges that are generally beyond the scope of the data producer’s expertise and bandwidth. Data archiving and distribution portals, such as the Earth System Grid Federation and the National Center for Atmospheric Research Data Archive, have greatly accelerated science through the provisioning of relevant climate data sets. However, more archival systems (and/or the expansion of existing portals) are needed to support the variety of products currently being used across the community. Opportunities exist for leveraging cloud services and/or server-side compute to potentially address these needs.

Avoiding redundancy and leveraging limited computational resources: Production of climate data products, particularly high-resolution products generated from process-based models, generally requires extensive computational resources and substantial human investments of time and effort to both run the models and archive the data. Facilitating better communication among data producers could identify needs that are addressable through coordinated simulations and make better use of existing computational resources. For example, better lines of communication could make an air quality modeler aware of community needs for wind power projections, and subsequently lead them to include high-frequency hub-height wind speeds as a model output. Additionally, the aforementioned identification of gaps among existing products could allow the community to identify high-priority simulations that have the broadest potential value.

Developing expert analysis and insights for data users: The choice of climate data products employed for decision-making is often based on existing research networks or data availability. In general, there is little guidance available to end-users on whether these products and their associated parent climate models meet a minimum standard of quality for their purposes. Community-developed and supported templates for metadata, which could include criteria for data documentation and licensing, along with requirements for publication of metric scores from an established and community-support evaluation protocol, and guidelines on best practices and/or pitfalls for parent climate model and data product averaging and weighting, would be helpful for informing decision-makers and building confidence in those data products. This information would further support data selection for widely-used, government-led community activities, such as the National Climate Assessment and National Nature Assessment. An increasing focus of the climate data community on "scientific co-production" has also highlighted the increasing need for researchers and end-users to work together to address relevant knowledge gaps, and suggests that efforts should be made to identify questions about climate data products of greatest importance for decision-makers.

Continuously improving observational (training) data products: High-quality observational data sets underpin any climate data product. Observational data products are constructed through various means, generally from meteorological station, airborne, or satellite observations or a combination thereof. The need for continuous improvement arises from the sheer number of unconstrained choices made to develop a product in terms of gridding and/or managing data outages, changes in measurement technology, instrument relocation, and other requirements to produce (with or without homogenization) long-term, high-temporal-and-spatial-resolution fields. In addition to improving these observational products, uncertainties around these products need to be quantified since they can translate to corresponding uncertainties in future impact projections.

Nurturing a cohesive regional climate data product community to address these challenges: The November 2023 Workshop was not the first workshop to address regional climate data issues. Many of the topics discussed echoed themes of previous workshops, but none of them resulted in a sustained, coordinated set of research activities to address long-standing, and more importantly, growing challenges with regional climate data and their connections to decision-making. The lack of a cohesive community to address the challenges discussed in the November 2023 Workshop was glaringly apparent. Workshop participants concluded that a follow-up workshop in 12-18 months would allow us to ascertain progress and plan for the future.

Research Needs for the Decision-Relevant Climate Data Community

The workshop concluded that substantial near-term progress could be made in addressing the eight challenges above, and laid out several potential deliverables that could also support longer-term improvements.

A community of practice: Conversations at the workshop highlighted the pressing need to ensure that lines of communication remain open between data producers, evaluators, and end-users. A community of practice, involving regular meetings and other means of facilitating communications between affiliated parties, would allow the climate data community to evolve to meet the ever-changing needs of this space. Beyond improving communication, we need to set forth a governance structure, a scope of activities, and incentives. These are critical for ensuring cohesion for the nascent community so that it can achieve and regularly measure its progress on the challenges identified in the workshop.

A common format for decision-relevant climate data: Early efforts by the Coupled Model Intercomparison Project (CMIP) led to the creation of a metadata standard and set of common variable names that would ensure interoperability of model data. Researchers have benefited greatly from this foresight, as most analysis tools and workflows can now be rapidly applied to model outputs, whether they be from Europe, Asia, or North America. Publicly-available tools such as the Climate Model Output Rewriter (CMOR) allow operational modeling centers to convert their native model outputs to data that conforms to a community standard. However, these practices have not been widely adopted by the regional climate modeling community, leading to workflows typically tailored to a particular data product. A common framework, decided upon by the climate data community, that specifies file format, metadata requirements and variable naming conventions would accelerate the usefulness of decision-relevant climate data.

A common framework for climate data product evaluation: Quantification of the performance of climate data products is an important step in ascertaining confidence in the data for decision support. With no commonly accepted standards for climate data evaluation, it is difficult to compare climate data products

.

and understand issues that may support or preclude their use. Consequently, there is a substantial and outstanding need for a community-developed framework for decision-relevant climate data product evaluation that leverages observation datasets and physical principles. Such a framework would identify and prioritize metrics, diagnostics, and other criteria relevant to the credibility of the data product. Providing accompanying expert guidance would assist in explaining observed differences between data and observations, and support the development of new strategies for climate data generation. This framework must also accommodate and navigate the differences inherent in the different types of downscaling and bias correction approaches.

Climate data cyberinfrastructure: Cyberinfrastructure to support the climate data community could include a maintained catalog of climate data products, disk space, and bandwidth to support archiving and provisioning of climate data and a computing platform for server-side analysis of climate data. Coordination among agencies could avoid redundant investments, ensure greater sharing of data, and allow users to avoid difficulties associated with accessing data through multiple platforms

Acknowledgments

This report is supported in part by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and was developed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Workshop funding and report writing was provided in part by a contract from the Strategic Environmental Research and Development Program (SERDP) under Project RC19-1391 and the Environmental Security Technology Certification Program (ESTCP) under Projects CR24-8399 and CR24-8916. The Workshop Scientific Steering Committee acknowledges Ellen Westfall of Noblis, Troy Cortez of LBNL, and Thor Swift of LBNL for organizing the workshop, hosting the workshop, and workshop photography, respectively.

Acronyms and Abbreviations

AI/ML	artificial intelligence/machine learning
ANL	Argonne National Laboratory
BMA	Bayesian Model Averaging
CAP	Climate Adaptation Partnerships Program
CASCs	Climate Adaptation Science Centers (USGS)
CF	Climate and Forecast
CMEC	Coordinated Model Evaluation Capabilities
CMIP	Coupled Model Intercomparison Project
CMIP5	Coupled Model Intercomparison Project – Phase 5
CMIP6	Coupled Model Intercomparison Project – Phase 6
CMOR	Climate Model Output Rewriter
CONUS	contiguous United States
COP	community of practice
CORDEX	Coordinated Regional Climate Downscaling Experiment
CRIS	Climate Resilience Information System
CUAHSI	Consortium of Universities for the Advancement of Hydrologic Science, Inc.
DECK	Diagnostic, Evaluation and Characterization of Klima
DOD	U.S. Department of Defense
DOE	U.S. Department of Energy
DRCDP	decision-relevant climate data products
E3SM	Energy Exascale Earth System Model
ECS	Equilibrium Climate Sensitivity
EPA	U.S. Environmental Protection Agency
EPRI	Electric Power Research Institute
ESGF	Earth System Grid Federation
ESM	Earth system model
ESTCP	Environmental Security Technology Certification Program
FAIR	findable, accessible, interoperable, and reusable
FARG	Federal Adaptation and Resilience Group
GCM	global climate model
GCRM	global cloud-resolving models
GLISA	Great Lakes Integrated Sciences and Assessments
GRL	Geophysical Research Letters
HDA	hierarchical data analyzer
HEALPix	Hierarchical Equal Area isoLatitude Pixelization
IGIM	Interagency Group on Integrative Modeling
IHTM	Integrated Hydro-Terrestrial Modeling
IPCC	Intergovernmental Panel on Climate Change
IDF	intensity-duration-frequency

JGR	Journal of Geophysical Research
JHM	Journal of Hydrometeorology
JPL	Jet Propulsion Laboratory (NASA)
LBNL	Lawrence Berkely National Laboratory
LLNL	Lawrence Livermore National Laboratory
LMRDD	learned multi-resolution dynamic downscaling
LOCA2	Localized Constructed Analogs, Version 2
NA-CORDEX	North American Coordinated Regional Climate Downscaling Experiment
NASA	National Aeronautics and Space Administration
NCA	National Climate Assessment
NCA5	Fifth National Climate Assessment
NCA6	Sixth National Climate Assessment
NCAR	National Center for Atmospheric Research
NCEI	National Centers for Environmental Information (NOAA)
NEX-GDDP-CMIP6	NASA Earth eXchange Global Daily Downscaled Projections
NOAA	National Oceanic and Atmospheric Administration
NOAA-CAPs	NOAA Climate Adaptation Programs
OCONUS	outside the continental United States
ORNL	Oak Ridge National Laboratory
OSTP	Office of Science and Technology Policy
PCAST	President's Council of Advisors on Science and Technology
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PDF	probability density function
PGW	Pseudo-Global Warming
PNNL	Pacific Northwest National Laboratory
RCM	regional climate model
RRM	regionally refined model
SCAN	Science for Climate Action Network
SCREAM	Simple Cloud-Resolving E3SM Atmosphere Model
SERDP	Strategic Environmental Research and Development Program
SR	super resolution
SST	sea surface temperature
STAR-ESDM	Seasonal Trends and Analysis of Residuals Empirical-Statistical Downscaling Model
SUNY	State University of New York
UCLA	University of California – Los Angeles
UHR	ultra-high-resolution
UIFL	Urban Integrated Field Laboratories (DOE)
USDA	U.S. Department of Agriculture
USGCRP	U.S. Global Change Research Program
USGS	U.S. Geological Survey
WCRP	World Climate Research Programme

Contents

Executive Summary iii
Challenges for the Decision-Relevant Climate Data Community iv
Research Needs for the Decision-Relevant Climate Data Community v
Acknowledgments iv
Acronyms and Abbreviations v
1.0 Context of the Workshop 1
1.1 Chapter Summary 1
1.2 Background 2
1.3 The Decision-Relevant Climate Data Product Landscape 4
1.4 Understanding Decision-Relevant Regional Climate Data Products 6
1.5 Workshop Structure 7
1.6 Report Structure 9
1.7 References 9
2.0 The State of Decision-Relevant Climate Data Production 11
2.1 Chapter Summary 11
2.2 Background 11
2.3 Data Production Methods 12
2.3.1 Statistical Downscaling 12
2.3.2 Dynamical Downscaling 13
2.3.3 AI/ML-Based Methods 15
2.3.4 Regionally Refined Global Modeling 15
2.3.5 Global Cloud-Resolving Models 16
2.3.6 Libraries of Short-Term Simulations 16
2.3.7 Common Gaps in Downscaling 17
2.3.8 Data Provisioning 17
2.4 References 19
3.0 The State of Knowledge on Users' Climate Information Needs 24
3.1 Chapter Summary 24
3.2 Background 24
3.3 Types of Climate Data Users 24
3.4 Types of Actionable Information 25
3.5 Types of Uses of Climate Information 26
3.6 Boundary Agencies Engaging with Climate Information Users 26
3.7 Progress in Meeting User Needs for Climate Information 26
3.7.1 Collaborative Approaches to Improve the Actionability of Climate Data 26
3.7.2 Examples of Collaborative Efforts 27

3.8 Remaining Gaps in Meeting User Needs 27
3.8.1 Gaps in Available Data and Tools 27
3.8.2 Lack of Guidance on Appropriate Use of Climate Data for Decision Applications 27
3.8.3 Lack of Credibility Evaluations from a Practical Standpoint 28
3.8.4 Gaps in User Engagement and Resources to Support it 29
3.9 Way Forward 29
3.10 References 30
4.0 The State of Benchmarking and Evaluation for Regional Climate Data Sets and Projections 32
4.1 Chapter Summary 32
4.2 Background 32
4.3 Standardized Evaluation of Decision-Relevant Climate Data 33
4.4 Inherited Biases from GCMs 34
4.5 Data Averaging and Weighting 34
4.6 Metrics and Diagnostics 35
4.7 Other Gaps and Challenges in DRCDP Evaluation 37
4.8 References 38
5.0 Technical Talks 40
5.1 Chapter Summary 40
5.2 Envisioning the Next Generation of U.S. Climate Predictions and Projections 40
5.3 Model Weighting Based on Equilibrium Climate Sensitivity 40
5.4 A Multi-Resolution Framework for Evaluating High-Resolution Climate Simulations 41
5.5 A Seamless Approach for Evaluating Climate Models across Spatial Scales/Evaluating Impact of Bias Correction on Downscaling Uncertainty 41
5.6 Assessing Physical Climate Risks: Challenges and Opportunities 41
5.7 Designing and Using Heterogeneous Ensembles of Climate Scenarios for Decision-Making 42
5.8 The North American Coordinated Regional Climate Downscaling Experiment (NA-CORDEX): Overview and Status 42
5.9 A Unified Framework for Evaluating the Risk of Extreme Temperature Events 43
5.10 Making Complex Data Actionable for Regional Decision-Making via the Cal-Adapt: Analytics Engine 43
6.0 Summary and Future Vision 44
6.1 Chapter Summary 44
6.2 Background 44
6.3 Summary of Identified Gaps 44
6.3.1 Gaps in Data Production 44
6.3.2 Gaps in Data Use 45
6.3.3 Gaps in Data Evaluation 45
6.3.4 Cross-Cutting Gaps: Observational Data 45

6.3.5 Cross-Cutting Gaps: Cyberinfrastructure 46
6.4 Parallels with the Coupled Model Intercomparison Project 46
6.5 Upcoming Challenges and Opportunities 46
6.6 Definition of a Community of Practice 47
6.7 References 49

Figures

1 The temporal and spatial scales of weather and climate phenomena that require decision-relevant climate projections, along with a depiction of the scales covered by CMIP6 and modern high-resolution ("High-Res") GCMs........ 5
2 In-person attendees of the DRCDP Workshop from November 14-16, 2023. 8
3 Breakout discussions at the November 14-16, 2023 DRCDP Workshop. 9
4 Production workflows for different decision-relevant climate data production methods. ESM stands for Earth system model, RCM stands for regional climate model, and RRM stands for regionally refined model. 12
5 A regionally refined model grid used for simulating climate over the contiguous United States. 16
6 Categorization of different climate data user groups developed for the Cal-Adapt Analytics Engine project and select examples. 25
7 Typology of actionable climate information derived from the iterative co-production engagements conducted in the DOE-funded HyperFACETS project. 26
8 Potential types of use or decisions by sector, management issue, and individual decision or use-case........ 25
9 Typology of uses of climate information derived from the iterative co-production engagements conducted in the DOE-funded HyperFACETS project. 25
10 A schematic representation of the evaluation workflow connecting data products to output supporting data producers and end-users. 33
11 A visualization of GCM performance across standard climatological metrics using a portrait plot. 33
12 Communication pathways that are needed and potentially addressed by a community of practice........ 48

1.0 Context of the Workshop

Lead Authors:

Daniel Feldman, Lawrence Berkeley National Laboratory, drfeldman@lbl.gov

Paul Ullrich, Lawrence Livermore National Laboratory, ullrich4@llnl.gov

Contributing Authors:

Moetasim Ashfaq, Oak Ridge National Laboratory, mashfaq@ornl.gov

Jared Bowden, North Carolina State University, jhbowden@ncsu.edu

Melissa Bukovsky, University of Wyoming, melissa.bukovsky@uwyo.edu

Ankur Dixit, Cornell University, ad933@cornell.edu

Keith Dixon, National Oceanic and Atmospheric Administration, keith.dixon@noaa.gov

Jeremy Giovando, U.S. Army Corps of Engineers, Jeremy.J.Giovando@usace.army.mil

Emiliano Gonzalez-Santin, Federal Emergency Management Agency, emiliano.gonzalezsantin@fema.dhs.gov

Ethan Gutmann, National Center for Atmospheric Research, gutmann@ucar.edu

Kripa Jagannathan, Lawrence Berkeley National Laboratory, KAJagannathan@lbl.gov

Andrew Jones, Lawrence Berkeley National Laboratory, adjones@lbl.gov

Sepideh Khajehei, NASA Ames Research Center, sepideh.khajehei@nasa.gov

Muge Komurcu, NASA Ames Research Center, muge.komurcu@nasa.gov

Rao Kotamarthi, Argonne National Laboratory, vrkotamarthi@anl.gov

David Lafferty, University of Illinois – Urbana-Champaign, davidcl2@illinois.edu

Hugo Lee, NASA Jet Propulsion Laboratory, huikyo.lee@jpl.nasa.gov

Ruby Leung, Pacific Northwest National Laboratory, ruby.leung@pnnl.gov

Fred Lipschultz, U.S. Global Change Research Program, flipschultz@usgcrp.gov

Jeremy Littell, U.S. Geological Survey, jlittell@usgs.gov

Elias Massoud, Oak Ridge National Laboratory, massoudec@ornl.gov

Rachel McCrary, National Center for Atmospheric Research, rmccrary@ucar.edu

Tanya Spero, U.S. Environmental Protection Agency, spero.tanya@epa.gov

Adam Terando, U.S. Geological Survey, aterando@usgs.gov

Nathalie Voisin, Pacific Northwest National Laboratory, nathalie.voisin@pnnl.gov

Casey Zuzak, Federal Emergency Management Agency, casey.zuzak@fema.dhs.gov

1.1 Chapter Summary

The need for localized, credible, authoritative, and accessible climate projections across federal, state, and local agencies, as well as large swaths of American society, is growing rapidly. Led by the U.S. Department of Energy Office of Science and the Strategic Environmental Research and Development Program (SERDP), several federal agencies were brought together to form a steering committee to organize a workshop to address this need. In conjunction with a team of scientists from national laboratories, academia and the broader community, the Workshop on Understanding Decision-Relevant Climate Data Products took place November 14-16, 2023, in Berkeley, California. The workshop covered the state-of-the-practice for downscaling and bias correcting climate models across the United States and using those results for adaptation and mitigation planning. The workshop focused on the science, organization and translation activities that are required to address the rising needs and interests of the nation for actionable climate information.

1.2 Background

A wide variety of federal, state, local, and private-sector decision-makers need data that is simultaneously salient, credible, authoritative, and accessible to pursue climate mitigation strategy, adaptation planning and vulnerability assessment. Decision-relevant climate data products (DRCDPs) are a crucial subset of the requisite data, but the need for an improved understanding of this data to improve their salience and credibility is a longstanding issue.

In the absence of guidance from a central authority, data users have relied on word-of-mouth or other inconsistent and ad hoc approaches for selecting climate data products. Recognizing this need, a number of federal initiatives and projects have been established to support these communities, including the DOE HyperFACETS project and the NOAA Climate Adaptation Partnerships Program (CAP). However, the pathways and mechanisms to support cross-agency coordination on climate information are still in their infancy. Developing guidance for data users also requires an improved understanding of the burgeoning landscape for climate information provision, especially from the private sector. Achieving that, in turn, requires a sustained assessment of confidence in the proliferating data sets based on an accepted framework for comparative evaluation of how robust they are for decision-making.

Box 1: Agency Perspectives

Numerous federal agencies are producing, analyzing, and using climate data products. Below are agency perspectives on the state-of-the-practice and needs for a workshop and climate information moving forward.

DOE Perspective: The US Department of Energy's mission encapsulates efforts to ensure America's security and prosperity by addressing its energy and environmental challenges. DOE's Office of Science works to support basic and applied climate science, including efforts focused on modeling and understanding extreme weather events and impacts (such as heat waves, atmospheric rivers, tropical cyclones, mesoscale convective systems, and other high-impact weather phenomena), subseasonal to decadal predictability, long-term projections and developing a deeper understanding of Earth system processes at all scales. The aforementioned research makes heavy use of hierarchical modeling, high-performance computing, and large data sets at decision-relevant scales. Further, development and provisioning of global, regional, and local climate data, particularly at the high spatial and temporal resolutions needed by decision-makers, directly supports this mission. DOE has a long background supporting global to regional climate model evaluation efforts, particularly through the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and has worked closely with regional stakeholders on questions related to climate data credibility and salience via the HyperFACETS project. This workshop provided an opportunity to further build on this exciting work.

SERDP/ESTCP Perspective: The rapidly advancing research in downscaled climate modeling has led to a growing number of tools, data and approaches that collectively serve the growing need for climate services and allow for climate change projections to meet the needs of many ancillary science disciplines. This interagency-sponsored workshop starts an important conversation that seeks understanding of this growing body of models, data, and applications, while preparing for continued improvements and coordination of future modeling.

NOAA Perspectives: NOAA is a leading provider of forward-looking environmental information for preparedness. For decades, the agency has invested in research to advance scientific understanding, modeling, prediction and projection of the causes and effects of changes in the climate system; and to advance services that help society plan and respond. Applying model-derived data for planning and decision-making is integral to NOAA's mission. To make model data useful for planning, thorough vetting and documentation is essential. Users need plain language summaries of the product's scope and intended use case(s) along with key details such as methods of production, calibration, and validation; known sources of error and uncertainty; peer review and provenance information; etc. For downscaled climate projections, we believe an extra burden of diligence is required to demonstrate and document model outputs' accuracy as compared to real-world observations for every parameter (variable), location and time of year. Additionally, the utility of particular products compared to other available data sets needs to be examined--which higher-resolution information is of prime interest to data users, process fidelity and information quality remain key. This will ensure that due diligence, including quality assessment and control measures, are an essential component of the development of new data and products by service providers. This workshop is an important step toward a needed common framework that all federal agencies and their partners can use to thoroughly vet and document their products and services.

EPA Perspective: EPA both develops and uses decision-relevant information. Although there are broad and global trends to climate change, the local effects are heterogeneous in space and through the seasons. Consequently, the methods for and abilities of communities to adapt are disproportionate across the Nation. Accordingly, the Nation requires scientifically sound *and localized* information about the potential changes to extreme weather events and regional and local climate to inform assessment, adaptation, and resilience activities to protect human health and the environment.

NASA Perspective: NASA's participation in this workshop aligns with its Earth Science to Action Strategy (NASA 2024), which emphasizes the translation of Earth science research into actionable information for societal benefit. As outlined in NASA's Climate Strategy (NASA, 2023), the agency provides precise, high-resolution observations and simulations that advance our understanding of both current and future climate. Through these efforts, NASA not only advances scientific discovery but also robustly supports policy-making and strategic decision-making across various sectors of government and industry.

FEMA Perspective: FEMA is both a producer and user of decision-relevant information. Emergency managers, hazard mitigation planners and community planners are responsible for taking actions to protect life, health, and property both now and as climate change impacts the current hazard landscape. The interagency workshop, and subsequent conversations, can help FEMA provide better science-backed guidance to its stakeholders on how to determine which data are relevant for specific decisions (i.e., can we use a particular future-oriented dataset to design flood-resilient structures? If the state of the science is not adequate for design, can we use a future-oriented dataset to help identify "low regret" adaptation options and test the sensitivity of those options against plausible climate impacts?). Furthermore, FEMA develops and makes available an array of natural hazard and risk information, and the Agency is exploring ways to incorporate climate change projections and/or modeling techniques in ways that make the information actionable and decision-relevant.

USGCRP Perspective: USGCRP, as part of its role in producing the National Climate Assessments and in taking on the effort to develop an architecture to provide information relevant to decisions for a changing climate (i.e., the Climate Resilience Information System, CRIS), has long supported the need to understand the landscape of climate projections. With the recent addition of climate services to the USGCRP remit, the goals of this workshop are important for providing the scientific underpinnings for informed decision-making across the nation.

1.3 The Decision-Relevant Climate Data Product Landscape

A multitude of global climate model (GCM) simulations have been undertaken within Coupled Model Intercomparison Project – Phase 6 (CMIP6) (Eyring et al. 2016). These simulations included both the meteorological conditions that were possible under historical climate, as well as those that could occur under a range of possible future scenarios. However, CMIP6 GCM data are typically available at spatial resolutions around 100km – generally too coarse for understanding climate and its potential future change on local to regional scales. Examples where GCM data may not be sufficient include: (1) meteorological and climatological conditions around local-scale land-surface features such as mountain peaks, urban areas, valleys and coastlines;
(2) calculations that require high temporal resolution due to strong nonlinearities or rapidly shifting conditions, such as renewable energy production or flooding; (3) extreme weather events, which often have outsized impacts at local scales, often occurring on scales of a few dozen kilometers or less and/or at short timescales; and (4) processes that are critical for climate adaptation and mitigation strategies for a given region where GCMs may have significant regional biases.

Figure 1 depicts the temporal and spatial scales associated with several important decision-relevant meteorological and climatological features, and highlights the limitation of CMIP6 and high-resolution GCM data for quantifying historical and future impacts. The time and space scales at which regional and local stakeholders are impacted by a changing climate are generally not included in global climate projections. To fill this gap, two basic strategies have emerged: first, the development of “ultra-high-resolution” (UHR) global climate models that operate at spatial scales near 4km, either globally or over a limited region; and second, the use of “downscaling” techniques to add fine-scale granularity to GCM data. The latter includes dynamical downscaling, which uses a regional climate model to simulate the regional weather using boundary conditions drawn from the GCM, and statistical downscaling, which uses empirical relationships between coarse and fine scales to interpolate GCM data to fine scales. Because they are relatively inexpensive, several statistically-downscaled products are available that map dozens of GCM simulations to spatial resolutions of 5km or finer, but rarely with temporal frequency higher than daily. Dynamically-downscaled and regionally-refined model products are increasingly available at these spatial scales and with sub-daily temporal resolution, but are more expensive to compute and so usually have fewer ensemble members. Because of their high computational cost, actionable UHR GCM products are still on the horizon, but are expected to play a greater role in the coming decade. More information on these techniques and their relative advantages and disadvantages can be found in chapter 2.

Because of the scientifically rooted history of GCM development and, by extension the CMIP project, the focus has primarily been on using models to build up the scientific community’s understanding of the Earth system. The considerations, priorities, and approaches for the use of the climate information by stakeholders are very distinct from the necessary academic research to improve Earth system understanding. For stakeholders, that understanding is a secondary priority: their primary need is for accurate projections with narrow uncertainty estimates, often for specific events, variables, or combinations of variables about which the stakeholder is concerned. Climate science for DRCDPs must advance skill for these types of projections to be relevant for stakeholder planning. To that end, GCM real or apparent biases over the historical record can seriously undermine the utility of GCM results for stakeholders. The natural variability of the Earth system also presents challenges to stakeholders since the near-term trajectory of the surface and atmosphere variables can be dominated by fluctuations that are difficult to characterize and predict.

Nevertheless, despite their relatively coarse resolution and the other challenges mentioned above, GCMs are also one of the scientific community's most powerful tools for evaluating future change. They are based on physical laws and representations of the whole Earth system and can therefore predict the response of the Earth system to imbalances caused by natural and/or anthropogenic sources (e.g., changing concentrations of atmospheric greenhouse gasses and aerosols and changing land-use and land cover). The importance of this capability for prediction cannot be overstated: the changes that the Earth system is experiencing, and will likely experience in the coming decades, are without precedent in recent history, so modeling the Earth system trajectory over this time period must rely on physical principles far more than historical observations and patterns.

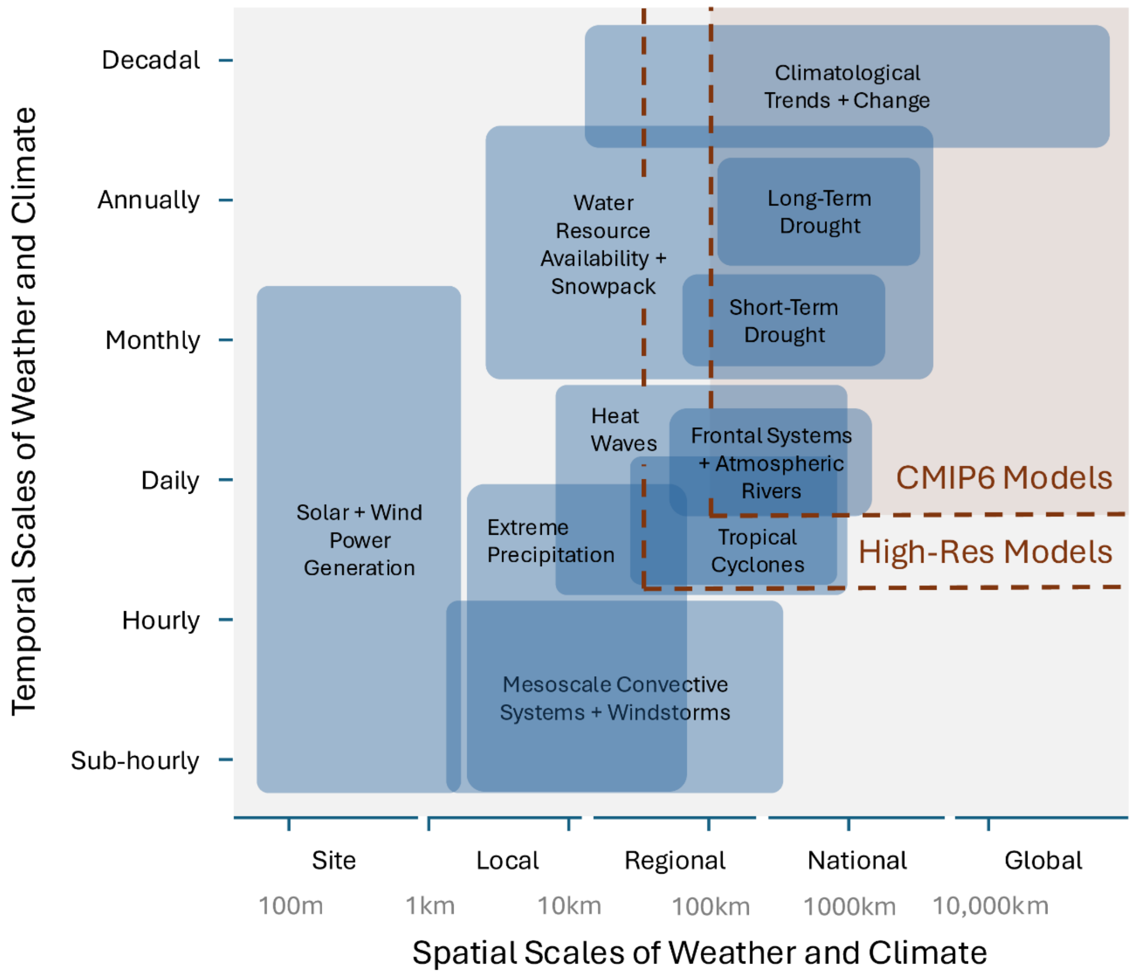


Figure 1. **The temporal and spatial scales of weather and climate phenomena that require decision-relevant climate projections, along with a depiction of the scales covered by CMIP6 and modern high-resolution ("High-Res") GCMs.**

As such, GCM data are almost always the starting point for developing higher-resolution climate data products. Directly simulating these scales with GCMs, while not impossible, is computationally prohibitive on a global scale for long periods of time. GCMs have also recently been built with support for regional refinement (localized high resolution), but at the time of this report, this solution is still an emerging technology. Consequently, decision-relevant data have been largely generated using post-hoc processing techniques. Specifically, numerous methods have been developed to "downscale" GCM data to the higher resolutions needed by a broad community of end-users (generally 1-25km). Downscaling involves techniques that introduce additional information on the relationships between processes at the GCM scale and those processes that impact the local scale, whether from historical patterns, physical modeling, or a combination of the two. A growing community of data producers have refined and applied these downscaling methods to generate decision-relevant climate data.

The desperate need for DRCDPs has meant rapid growth in the number of such data products coming online in recent years, particularly over the contiguous United States. Already the number of such products has prompted confusion from data users, who must decide among available options to focus their limited resources. Little public guidance is available from experts in the community, and almost none of it covers the broad space of available products. As a result, end-users have often been left to blindly navigate a space colloquially referred to as the "Wild West" of decision-relevant climate data products.

1.4 Understanding Decision-Relevant Regional Climate Data Products

Motivated by the need for more effective coordination across agencies and in the broader climate data community, parallel conversations on the best path forward occurred among the Interagency Group on Integrative Modeling (IGIM), the Federal Adaptation and Resilience Group (FARG), elsewhere in USGCRP and within federal agencies. The idea of a workshop arose in early 2023, with the goal of bringing together representatives from the entire climate data space, including data producers, analysts, end-users, agency representatives and scientists, to map out efforts currently underway and identify gaps and challenges limiting future progress. In large part, the premise of this coordinated effort was that a decision-relevant climate data community could achieve much more together than the sum of its parts.

Further motivation for the workshop came from the recent experience of selecting and evaluating climate projections for use in the Fifth National Climate Assessment (NCA5; USGCRP 2023). Among the required attributes were publicly available methodology, technical documentation, algorithms, and source code for the downscaling models, as well as publication in peer-reviewed scientific journals. Besides meeting FAIR principles (findability, accessibility, interoperability, and reusability; e.g., Wilkinson et al. 2016) and the Foundations for Evidence-Based Policymaking Act (2018) requirements, recommended attributes included providing variables relevant to climate impacts at multiple spatial/temporal scales and appropriate spatial and temporal resolutions for decision-making. Towards the end of the NCA5 process, funding from the USDOE supported an initial comparison of the two selected datasets (Ullrich 2023). The evaluation of data products for NCA5 was initiated only towards the end of the NCA5 process since there was no procedure in place for such efforts.

During the NCA5 process, it was already clear that there will be an expanding universe of projections based on data from the sixth phase of the Coupled Model Intercomparison Project (CMIP6) that could inform the Sixth National Climate Assessment (NCA6), as well as serve a vast array of adaptation needs across the nation. Inclusion in NCAs implies confidence that the dataset is sufficiently robust for use in decision-making; however, the process for NCA5 focused largely on availability for the NCA5 timeline, with the comparison limited to just two statistical data sets that were available prior to the report release in November 2023.

For the NCA6, as well as for USGCRP's new emphasis on delivering climate services (herein defined as the provision and use of climate data, information, and knowledge to assist decision-making), there is a clear need for an approach requiring adherence to FAIR standards and the Foundations for Evidence-Based Policymaking Act (2018), and also for assessing confidence for still-to-be determined metrics of robustness or benchmarks of a downscaled dataset for decision-making. Developing comparable information about each data set will be greatly facilitated by improving and standardizing the underlying metadata across the diverse and rapidly expanding data sets available to support the vastly increased and consequential suite of user needs.

Beyond NCA6, the proliferation of new data will provide better constraints on future change, but, in the absence of an intervention, are expected to cause greater confusion among decision-makers, scientists, translators and end-users about how to identify and tailor the most appropriate climate information for the myriad of applications across every community, business, or agency in the nation. Improving our common knowledge base about the available downscaled products is an important step towards an effective decision support system. New, community-based approaches are needed to advance the capacity of climate-sensitive decision-makers to evaluate the appropriate use of climate projections to make informed decisions on how best to prepare for, and adapt to, climate change.

1.5 Workshop Structure

The workshop was designed within the context of the above background – to understand the state of the nation's decision-relevant regional climate projections and carry that understanding forward, so as to enable the development of guidelines for the usage and evaluation of such projections. To achieve this goal, the workshop aimed to share knowledge between producers, users, and evaluators of downscaled data, establish common ground, and begin to build a community of practice. The ultimate objective was development of guidelines for production, evaluation, and use of high-resolution regional climate projections of impact-relevant variables.

Participation came from most major U.S. federal agencies and their partners who are involved in the production and dissemination of regional climate data products. These include DOE, NOAA, NASA, EPA, FEMA, USGS, USBR, and DOD's Strategic Environmental Research and Development Program and Environmental Security Technology Certification Program (SERDP and ESTCP). Box 1 contains perspectives from each of the participating agencies on their specific needs and interests in advancing the use of, and confidence in, climate data products. Dozens of in-person attendees visited Berkeley, California from November 14-16, 2023 and were joined by dozens more virtual attendees. Figure 2 shows a group photo of the in-person attendees.

Figure 2. In-person attendees of the DRCDP Workshop from November 14-16, 2023. © The Regents of the University of California, Lawrence Berkeley National Laboratory.

Numerous approaches for generating regional climate data were discussed, including statistical downscaling, dynamical downscaling, hybrid downscaling, regionally refined global modeling and artificial intelligence. This effort made great strides in sharing knowledge, establishing common ground, and moving towards the development of a community of practice around decision-relevant data. As seen in Figure 3, the workshop was an effective opportunity for networking and conversations among the leaders in this field.

The workshop was organized into four sessions that focused on (1) data production; (2) data use; (3) data evaluation; and (4) emerging topics. The session on data production featured ten talks from a variety of data producers, representing multiple federal agencies and academic research groups, followed by breakout sessions that sought to frame the needs of a community of practice. The session on data use featured two panels, each with four panelists presenting brief talks on topics related to how they employ climate data and their perceptions of gaps among existing data products. The session on data evaluation again featured two panels, each with four panelists presenting brief talks related to ascertaining the credibility of climate data. The final session on emerging topics featured 11 technically oriented talks on topics related to climate data, including bias correction, model weighting, ensembles, and performance across scales. For each of the first three sessions, there were accompanying breakout discussions which featured a mix of participants framed around key questions related to that session, in the context of the broader workshop theme.

Figure 3. Breakout discussions at the November 14-16, 2023 DRCDP Workshop. © The Regents of the University of California, Lawrence Berkeley National Laboratory.

1.6 Report Structure

This report summarizes the discussions from the workshop and is roughly structured to cover the workshop's three core themes: chapter 2 addresses challenges and gaps in decision-relevant climate data production, chapter 3 covers needs from climate data end-users, chapter 4 focuses on climate data evaluation, chapter 5 covers ongoing and future research needs and chapter 6 summarizes identified gaps and suggests a strategy for the development of a community of practice around decision-relevant climate data. This report is reflective and not exhaustive: it aims to present a focused discourse relevant to climate scientists and stakeholders concerning the state of the science and existing gaps, rather than encompass a comprehensive review or encapsulate all viewpoints.

1.7 References

Foundations for Evidence-Based Policymaking Act. 2018. H.R. 4174, 115th Congress. https://www.congress.gov/bill/115th-congress/house-bill/4174

NASA. 2023. Advancing NASA's Climate Strategy, https://www.nasa.gov/wp-content/uploads/2023/04/advancing-nasas-climate-strategy-2023.pdf

NASA. 2024. Earth Science to Action, https://science.nasa.gov/earth-science/earth-science-to-action/

Ullrich, PA. 2023. Validation of LOCA2 and STAR-ESDM Statistically Downscaled Products. LLNL-TR-856094; 1085545, https://doi.org/10.2172/2202926

USGCRP. 2023. Fifth National Climate Assessment. AR Crimmins, CW Avery, DR Easterling, KE Kunkel, BC Stewart, and TK Maycock, Eds. U.S. Global Change Research Program, Washington, D.C., USA, https://doi.org/10.7930/NCA5.2023

Wilkinson, M, M Dumontier, I Aalbersberg, G Appleton, M Axton, A Baak, N Blomberg, J-W Boiten, L Bonino da Silva Santos, PE Bourne, J Bouwman, AJ Brookes, T Clark, M Crosas, I Dillo, O Dumon, S Edmunds, CT Evelo, R Finkers, A Gonzalez-Beltran, AJG Gray, P Groth, C Goble, JS Grethe, J Heringa, PAC ’t Hoen, R Hooft, T Kuhn, R Kok, Joost Kok, SJ Lusher, ME Martone, A Mons, AL Packer, B Persson, P Rocca-Serra, M Roos, R van Schaik, S-A Sansone, E Schultes, T Sengstag, T Slater, G Strawn, MA Swertz, M Thompson, J van der Lei, E van Mulligen, J Velterop, A Waagmeester, P Wittenburg, K Wolstencroft, J Zhao, and B Mons. 2016. “The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific Data* 3: 160018, https://doi.org/10.1038/sdata.2016.18

2.0 The State of Decision-Relevant Climate Data Production

Lead Author:
Rao Kotamarthi, Argonne National Laboratory, vrkotamarthi@anl.gov

Contributing Authors:
Moetasim Ashfaq, Oak Ridge National Laboratory, mashfaq@ornl.gov
Jared Bowden, North Carolina State University, jhbowden@ncsu.edu
Melissa Bukovsky, University of Wyoming, melissa.bukovsky@uwyo.edu
Dan Cayan, University of California – San Diego, dcayan@ucsd.edu
Ryan Currier, National Oceanic and Atmospheric Administration, william.r.currier@noaa.gov
Ankur Dixit, Cornell University, ad933@cornell.edu
Katrina Fandrich, University of Albany, kfandrich@albany.edu
Daniel Feldman, Lawrence Berkeley National Laboratory, drfeldman@lbl.gov
Ethan Gutmann, National Center for Atmospheric Research, gutmann@ucar.edu
Muge Komurcu, NASA Ames Research Center, muge.komurcu@nasa.gov
Ruby Leung, Pacific Northwest National Laboratory, ruby.leung@pnnl.gov
Jeremy Littell, U.S. Geological Survey, jlittell@usgs.gov
Rachel McCrary, National Center for Atmospheric Research, rmccrary@ucar.edu
David W. Pierce, University of California - San Diego, dpierce@ucsd.edu
Stefan Rahimi, University of Wyoming, srahimi@uwyo.edu
Adam Terando, U.S. Geological Survey, aterando@usgs.gov
Oliver Elison Timm, University of Albany, oelisontimm@albany.edu
Tanya Spero, U.S. Environmental Protection Agency, spero.tanya@epa.gov
Paul Ullrich, Lawrence Livermore National Laboratory, ullrich4@llnl.gov

2.1 Chapter Summary

Because the spatial resolution of global climate model outputs is typically around 100km, a wide variety of downscaling techniques have been developed to supplement climate model projections with information on the processes that impact climate locally. These techniques can take quite different approaches to downscaling. This chapter focuses on the state-of-the-practice for statistical and dynamical downscaling and touches on some of the strengths and weaknesses of those techniques. New frontiers for climate model downscaling research and applications are also touched upon, including work underway on artificial intelligence and machine learning (AI/ML)-based methods.

2.2 Background

The need for spatial downscaling of climate model outputs for assessing the impacts at regional and local scales was noted very early in the IPCC assessment cycle (Gates 1985). Downscaling methods grew from simple spatial disaggregation (e.g., Wood et al. 2002) of GCM outputs to more sophisticated statistical (Wilby et al. 1998, Pierce et al 2014, Pierce and Cayan 2016, Gutmann et al. 2022) and, increasingly, dynamical downscaling approaches (Bowden et al. 2012, Otte et al. 2012, Mearns et al. 2014, Prein et al. 2017, Komurcu et al. 2018, Rahimi et al 2024a, Wang and Kotamarthi 2015, Rasmussen et al. 2014, etc.). Hybrid methods that combine statistical and dynamical downscaling have also been developed. Using AI/ML models to downscale climate models has rapidly developed over the past few years and can be expected to become a key method for downscaling in the near future (e.g., Huntingford et al. 2019, Hobeichi et al. 2023). Simultaneously, approaches that operate climate models at increasingly finer grid spacing globally or with grid refinements (Fox-Rabinovitz et al. 2005, Zarzycki et al. 2015) over regions of interest from within a global model are becoming more available and viable for performing multi-decadal and multi-ensemble simulations. Figure 4 depicts the general workflows for each of these methods, which highlights both similarities and differences among these methods.

A consistent theme of the workshop discussions was that the use of multiple DRCDPs could better represent the broad range of scientific understanding of climate-sensitive physical processes that can impact planning, and that a single DRCDP could underestimate or overestimate risk. Conversations at the workshop emphasized that there is no single approach that should be the basis for informing decisions, but that a rich ecosystem of methods can provide multiple lines of evidence to constrain future uncertainties. Notably, each of these methods does not necessarily exist in isolation, and new techniques are continually being developed that hybridize these methodologies, adopting features from more than one approach.

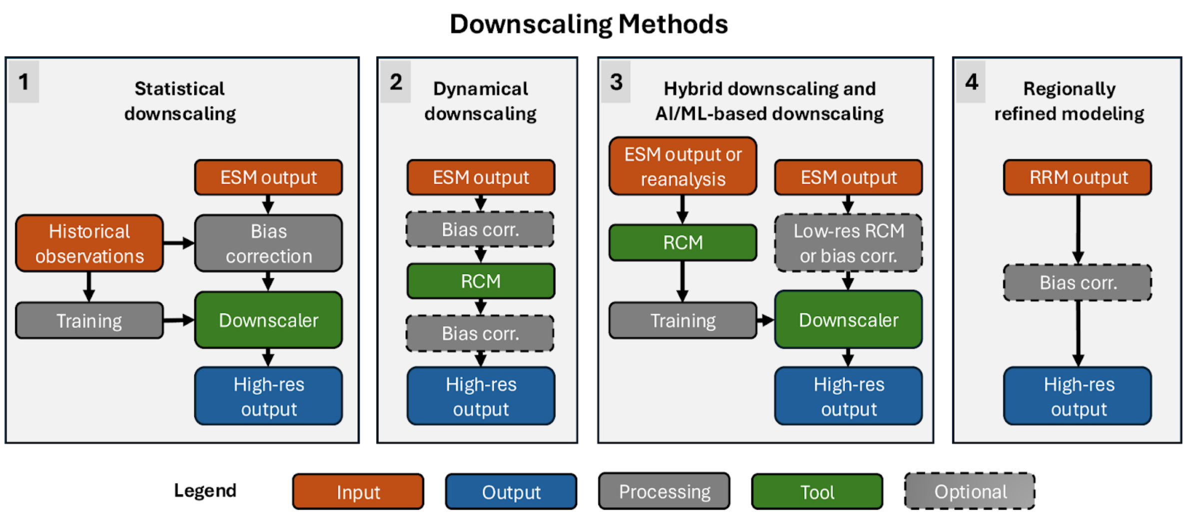


Figure 4. Production workflows for different decision-relevant climate data production methods. ESM stands for Earth system model, RCM stands for regional climate model, and RRM stands for regionally refined model.

2.3 Data Production Methods

2.3.1 Statistical Downscaling

Statistical downscaling methods include a wide range of approaches and are by far the most widely used for decision-relevant climate data products. The relative ease of implementing these methods with GCM outputs and the ability to produce large ensembles of bias corrected results has made them popular with end-users. Their fidelity relative to historical observations can be demonstrated, and they are more realistic at local scales than GCM outputs. At the workshop, presentations were given on statistical downscaling products with a mix of national and global coverage, including the Localized Constructed Analogs, Version 2 (LOCA2) (Pierce et al, 2023), the Seasonal Trends and Analysis of Residuals Empirical-Statistical Downscaling Model (STAR-ESDM) (Hayhoe et al, 2023) and NASA Earth eXchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6) (Thrasher et al, 2022).

The primary challenges for this method are (1) extending the downscaling beyond surface air temperature and precipitation to a larger number of variables due to the paucity of appropriate local weather data and (2) preserving the dynamical and/or thermodynamic consistency of variables and maintaining known physical relationships between these variables when doing so.

Statistical methods trained on historical data inherently incorporate assumptions of stationarity: they assume the spatial patterns of the past remain the same into the future, despite evidence, theory, and a fundamental understanding of Earth system dynamics that suggest nonstationarity, for example, shifts in storm tracks (Yin 2005, Bengtsson et al. 2006, Ulbrich et al. 2008, O'Gorman 2010), sharpening of precipitation (Chen et al., 2023), and other dynamical and thermodynamical changes in the Earth system are expected in the coming decades due to the global hydrological response of the Earth system to warming (Jeevanjee and Romps, 2018). Furthermore, presently available statistical methods struggle with variables that require a "memory" of past conditions, such as snowpack and soil moisture, and cannot capture related feedbacks (e.g., higher near-surface relative humidity when soil moisture and subsequent evapotranspiration is high).

For statistical downscaling, the choice of gridded historical products, application of bias correction and details of the downscaling algorithms affect uncertainty in the results (Ullrich 2023, Lafferty and Sriver 2023). An understanding of how to quantify these uncertainties and communicate this information remains a conspicuous research gap.

2.3.2 Dynamical Downscaling

Under dynamical downscaling, GCM projections are used as initial and lateral boundary conditions to drive higher-resolution RCMs. The ability of dynamical downscaling to better capture non-stationary changes in future climates is valuable for decision-makers, and distinct from what statistical methods provide.

Like GCMs, RCMs have land surface and atmospheric components and use established physical parameterizations based on theory, observations, and modeling. Internal consistency among variables in RCMs is then achieved via these model components and parameterizations.

At the workshop, several groups (e.g., Argonne National Laboratory (ANL), Oak Ridge National Laboratory (ORNL), Environmental Protection Agency (EPA), National Center for Atmospheric Research (NCAR), University of California-Los Angeles (UCLA), State University of New York at Albany (SUNY Albany) and Pacific Northwest National Laboratory (PNNL)) spoke about unprecedented and ambitious work underway to dynamically downscale CMIP6 GCMs. These efforts have made major advances in understanding the physical process representations and model configurations needed for high-quality representation of local climate processes and relevant nonstationarities. However, each effort seeks to investigate unique questions and geographies, while also being a major undertaking in terms of person-power, computational resources, and data science.

Despite its fundamental strengths and recent advances, dynamical downscaling continues to be computationally limited to a few GCMs, especially compared to statistical downscaling. Typically this means selecting one to three GCMs that are representative of the physical processes affecting the region of interest, which provides an abridged view of the uncertainty space afforded by including more GCMs. Similarly, only a limited number of socio-economic scenarios can be explored in the analysis, again limiting the decision maker's choices, and further narrowing the range of plausible futures.

At the same time, many groups are pursuing dynamical downscaling. The sheer number of disconnected efforts to dynamically downscale GCMs at the continental, national, and sub-national scales signifies opportunities for the community: for instance, the pooling of computational resources to produce larger and more coordinated dynamically downscaled GCM ensembles across the contiguous United States (CONUS) and outside the continental United States (OCONUS) using different RCMs. One option is to capture the overlapping downscaled GCM simulations across the western U.S. for MPI-ESM1-2-HR from UCLA and ORNL, albeit with different RCMs. Given that 22 GCMs were downscaled for CMIP5 (CORDEX-SAT, 2020), it is reasonable to suggest that at least as many GCMs may be dynamically downscaled in CMIP7. Additional options are included in the section 'Common gaps in downscaling'.

In addition to coordination, several gaps related to improving dynamical downscaling were discussed at the workshop:

Bias correction: A critical challenge is the need for bias corrected model outputs to address both RCM biases and the biases they inherited from GCMs. There is an ongoing debate on whether bias correction should be applied to inputs of an RCM, outputs of an RCM, or neither. Several authors have found that bias correcting the GCM input fields to the dynamic downscaling models at continental scale does not appear to reduce the simulated bias on larger regional and continental scales (Xu and Yang 2015, Wang and Kotamarthi 2015). Recently, however, Rahimi et al. (2024b), Risser et al. (2024), and others have also been exploring the impacts of pre-downscaling bias correction and found that it does lead to greater skill in the RCM orographic precipitation, snowpack, and temperature simulation. An additional consideration is whether the bias correction should only target averages, or if producers should use a more complicated bias correction of the GCM boundary conditions. Post-simulation bias correction of RCMs has an extensive history, particularly in the hydrology community, and various methods have been explored (Maraun 2013, Adeyeri et al. 2020, Yang et al. 2015, Wilcke et al. 2013, Francois et al. 2020). This form of bias correction is often necessary because impact models (e.g., flooding, wildfires, energy, health) are often highly sensitive to meteorological inputs. Preservation of the correlation between variables that are generated by the downscaling and the time series of the distributions, while not altering the tails of the distributions, are major challenges for post-RCM bias correction that are still being investigated.

Pseudo-Global Warming (PGW): The PGW method, which applies GCM-derived climate deltas to a historical reanalysis before dynamical downscaling, allows one to estimate a high-resolution change signal in future climates. This approach was used recently to build a long-running high-resolution climate data product over the contiguous United States that represented the historical period and eight possible futures (Jones et al. 2023), and recent work has compared it to direct downscaling of GCMs (Hall et al. 2024). Examples of questions surrounding PGW's applicability include the following. First, is the assumption of time-invariant natural variability justifiable? For example, for intense heat waves, strong land-atmospheric coupling in a warmer world may lend itself to evapotranspiration reductions and sensible heat flux increases. This may lead to deepening of the 'heat dome' and an amplified heat wave, compared to expectations from the GCM warming delta. Second, the PGW method is sometimes applied using GCM ensemble deltas; however, given the unique thermodynamic and hydrologic sensitivities of each GCM, does an ensemble developed with this approach represent an actual ensemble of possible futures with acceptable levels of uncertainty incurred from this approach? Or, must unique PGW experiments be conducted for individual GCMs before average regionalized change signals are computed and examined?

Convection-permitting scale downscaling/regional climate modeling: Convection permitting regional climate modeling (with horizontal resolutions grid spacing $\lesssim$ 4km), was previously untenable because of its heavy computational and storage cost, but is becoming increasingly possible because of significant software and hardware advancements. This method has the potential to improve the representation of mean and extreme values of climate variables (Prein et al. 2017, Komurcu et al. 2018, Akinsanola et al. 2024), particularly because increased spatial and temporal resolution allows for more detailed interactions between land surface, boundary layer and cloud processes. As computational expense continues to decline, this methodology may produce simulations with process fidelity and consistency among variables to enable comparisons between models and across ensembles, and the exploration of a wider range of what-if scenarios.

Hybrid downscaling: There is emerging interest in using dynamical downscaling products, instead of gridded observations, as training data for statistical downscaling. This technique would greatly expand the number of variables available to statistical models and enable higher temporal resolution. It would also allow those models to be trained on both historical and future data, which should help to alleviate issues with stationarity. Only recently have sufficiently long duration simulations come online that could provide enough training data, and bias correction remains necessary to correct biases from the GCM and RCM.

Investigating parameterization stationarity: Many of the physical parameterizations used by both RCMs and their parent GCMs are based on (typically a small number of) empirical observations. Although substantial work has gone into generalizing those parameterizations to work well across a variety of geographies, it is far more difficult to demonstrate that parameterizations are valid outside the directly observed time period (e.g., Baumberger et al. 2017). Namely, additional research is needed to ensure that those parameterizations are climate-aware and are not inadvertently introducing stationarity into the non-stationary processes that dynamical downscaling intends to capture.

2.3.3 AI/ML-Based Methods

AI/ML has made extensive strides in the past decade and has emerged as a new approach for a wide range of scientific applications, including downscaling and bias correction. Although not yet operationalized, among those deep learning-based models that have shown success to date are super-resolution (SR) methods and learned multi-resolution dynamic downscaling (LMRDD) methods. SR methods use a high-resolution data set that is upscaled to a coarse resolution (e.g., the resolution of the GCM to be downscaled) to build a AI/ML model. These methods and variations have now been used to downscale wind (Stengel et al. 2020) and precipitation (Geiss and Hardin 2020). Conversely, the multi-resolution LMRDD method uses the GCM model output and its dynamically downscaled output (without any upscaling) to train the LMRDD model (Wang et al. 2021). As with statistical downscaling methods, SR models are limited by the availability of high-resolution observational data sets and suboptimal accuracy of coarse resolution modeled precipitation due to the physics parameterizations at these scales. LMRDD models are limited by the availability of dynamically downscaled data sets, which are comparatively rare. However, once developed, these models are significantly less computationally expensive than dynamical methods, so they can be employed to downscale an entire set of GCM simulations, scenarios, and time slices. A recent approach has been to use diffusion based models for downscaling (Ling et al. 2024).

2.3.4 Regionally Refined Global Modeling

RRMs are GCMs that have a non-uniform grid with high resolution over a particular geographical region, thus avoiding the need for a secondary RCM driven by GCM output and permitting two-way coupling between the high-resolution and global domains (e.g., Zarzycki et al. 2015, Tang et al. 2020). RRMs can also be employed as RCMs, for instance through nudging of the coarse region towards some reference data. However, only a few operational GCMs have support for regional refinement, given the algorithmic complexity usually necessitated by this approach. To date, RRMs have been more widely employed for modeling on weather time scales than climate time scales, but have demonstrated success in simulating several types of weather phenomena (e.g., Liu et al. 2023). A depiction of a regionally-refined mesh with coverage of the CONUS is given in Figure 5.

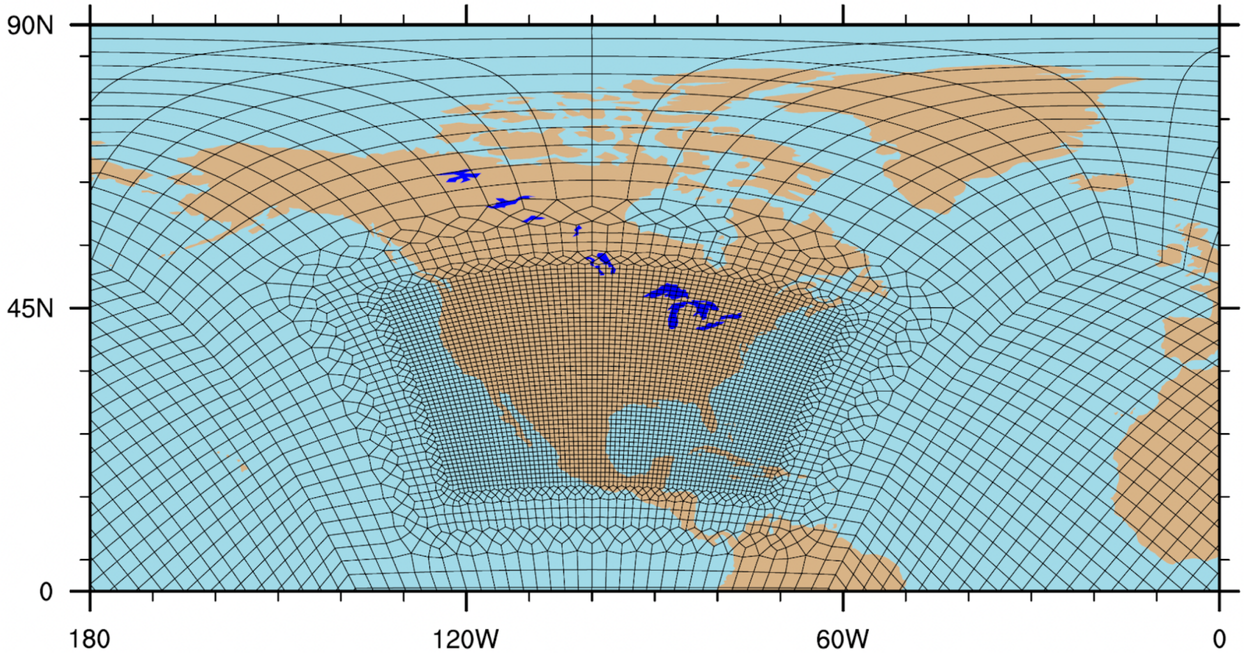


Figure 5. **A regionally refined model grid used for simulating climate over the contiguous United States. A higher density of grid cells provides more resolution and targets computational resources at the region of interest.**

2.3.5 Global Cloud-Resolving Models

Global cloud-resolving models (GCRMs) refer to GCMs that can explicitly resolve convective processes, thus avoiding the need for a convective parameterization (e.g., Donahue et al. 2024). It is expected that explicitly resolved dynamics will avoid persistent issues with parameterized convection (Molinari and Dudek 1992, Rio et al. 2019) and improve the simulation of sharp meteorological gradients (e.g., in extreme weather events). With grid spacing of the order of 4km or less, these models can only run on large supercomputing systems, and to date have only produced simulations of a few simulated years. Nonetheless, with exponential growth in computing power, there is an expectation that these models will be increasingly employed for modeling of the climate system in the next decade. While GCRMs still do not capture the finest scales of relevance in the atmosphere able to impact local climate (e.g., large eddy scales), they are starting to become an important benchmark for downscaling, particularly in regions that lack reliable observational data (e.g., OCONUS). GCRMs can also explore the sensitivity of local projections to the simulation of weather and climate processes at high-resolution across the Earth.

2.3.6 Libraries of Short-Term Simulations

Although the discussed data production efforts have primarily focused on large ensembles of long-term climate simulations, efforts could also focus on large ensembles of extreme events, performed over a shorter time window and thus permitting higher spatial resolution (e.g., Huang et al. 2020). These ensembles could better target computational resources and address end-user needs. Events could come from the historical record, and simulated using RCMs or RRMs, potentially in combination with the PGW methodology. They could also come from GCM large ensembles, leading to a larger sample of synthetic events or events beyond what has occurred in the historical record. Such ensembles naturally yield scenarios for assessing and quantifying impacts.

2.3.7 Common Gaps in Downscaling

Geographic Coverage: At the workshop, there was significant discussion about the lack of high-quality high-resolution climate data products for domains outside of CONUS. In particular, the U.S. Virgin Islands, Puerto Rico, Guam, Hawaii, and Alaska (a region collectively referred to as OCONUS) are all limited in both observations and simulated climate data products. High-resolution data are particularly relevant for the U.S. islands because local climate can vary significantly over the width of the island and differs from the surrounding ocean, meaning that presently-available global statistically downscaled products (typically available at 0.25 degrees) are still too coarse. The relative sparsity of the observing network in these regions and the need for high spatial resolution suggests a need for new data generated using dynamical downscaling (e.g., Fandrich et al. 2022, Mizukami et al. 2023). However, with few observations, validating climate data products is also difficult, likely necessitating supplementing in situ observations with satellite data. Beyond the U.S. and Europe, similar issues are also present, particularly among countries that don't have strong operational infrastructure for observing and simulating weather. Differences in the quality or source of observational data has also produced curious artifacts among presently available climate datasets; for example, many data products end abruptly at the U.S. border, even when these excluded regions are part of CONUS-relevant watersheds.

Availability of high-temporal resolution data: DRCDPs with sub-daily frequency are largely unavailable at present. Extreme weather events necessitate high-temporal resolution data, although the precise interval depends on the specific use case. For extreme storms that could lead to flooding, hourly data is desirable. Meteorology also significantly influences resource adequacy for energy supply and demand, particularly renewable resources. For example, estimates of wind power capacity factors, which are proportional to the cube of wind speed, are very sensitive to short-term variation of winds and so reasonably accurate calculation of power production requires at least hourly wind data. Going further, a real-time (sub-hourly) meteorological data set is necessary for an efficient electrical grid stabilization and management system, so as to characterize the magnitude of risk and variability, aiding in effective capacity planning and optimal scheduling (Fu et al. 2024).

Quantified uncertainties among data products: To ensure comparability and reliability across products, common techniques for quantifying uncertainty across climate datasets are needed. For example, while there are established methods for developing measures of uncertainty from a set of models and ensemble members, the selection of the ensemble size and the quality of the models used in generating the ensemble affect the outcome of this calculation.

Production of secondary impact variables: Risk and impact estimates that are of interest to stakeholders (e.g., drought, wildfires, inland and coastal floods) require additional processing and climate variables beyond those needed for assessing heat wave frequencies and precipitation intensities. Development of robust estimates of impact frequencies, intensities and duration and their associated uncertainties depends on having a large ensemble of model simulation outputs. These ensembles allow for more robust calculation of the return periods that are usually dictated by stakeholder needs. Additional downstream models are further needed for estimating these impacts (e.g., inland hydrology models).

2.3.8 Data Provisioning

Data provisioning refers to the distribution of data, along with accompanying details on how they were produced and how they should be used. Ideally, data provisioning efforts should aspire to FAIR principles (Wilkinson et al. 2016). Support for data producers to address FAIR principles remains a major challenge in the climate data community, particularly in light of the issues identified here.

Lack of downscaled data standards: Lack of consistency among climate data products can create confusion among data users and prevent interoperability of data products across workflows. Consistency here refers to both documentation of the data products and the data files themselves. Consistent documentation should include producer-developed descriptions of the data set's characteristics (resolution, geographic coverage, time period, etc.), how it was generated and how it should be used. At present, such information is either only available from direct engagement with data producers or is inconsistently presented in technical documentation and websites. Consistent data requires common formats, metadata, variable naming, and units, even among gridded observational products. To an extent, the Coordinated Regional Climate Downscaling Experiment (CORDEX) has been effective at spearheading standardization in the downscaling community (McGinnis and Mearns 2021). Nonetheless, additional work is needed to achieve widespread adoption of community standards for robustness.

Limited computational resources: Computational cost is high for long-period (i.e., 20 years but ideally 30) dynamically downscaled projections, especially when using convection-permitting ($\lesssim$ 4km) modeling. There is a need to recognize that not all groups that conduct dynamical downscaling have access to computing resources to support "sufficiently long" simulations or a "sufficiently large" ensemble or "sufficiently fine" resolution. Significant costs are associated with running the model, storing the data, post-processing, and dissemination. Also, end-users may want further processing, e.g., to set thresholds on resulting fields for specific risk questions, but these may not be possible to provide without rerunning the model. This, and similar interactive analysis capabilities, would benefit many end-users and relieve the pressure on end-users for finding computational resources to perform such calculations.

Data access: Data must be accessible by decision-makers if those data are going to support needs for resilience and adaptation planning. Whereas GCM output is generally contributed to a common repository, such as that maintained by the Earth System Grid Federation and replicated by other groups worldwide, downscaled data do not have a unified portal or interface for distribution. Reasons include lack of a common domain (i.e., subset of the globe), lack of a common spatial and temporal resolution, lack of a common suite of output variables and size of the available output. In addition, differences in downscaling methods, philosophies and other scientific decisions make different instantiations of downscaled data more appropriate for some use cases than others. Furthermore, differences and availability of computing infrastructure and resources among the groups that develop the data (including computer security limitations, disk availability, documentation, and staff limitations) can inhibit sharing downscaled data with external users.

Spatial disaggregation: Local decision-makers often ask for very high-resolution data from climate models/data products (<1km) that does not exist and would be very costly to produce and distribute. However, upon engaging with these groups it is sometimes clear that the data they presently use is highly uncertain. Consequently, it is important to understand for what spatial scales and purposes these high-resolution datasets actually provide benefit over coarser resolution products. As noted in Ullrich (2023), although many climate data products provide information on a higher-resolution grid, the credible resolution of those data sets are likely to be much coarser. In fact, in some circumstances it may be better to use coarser-resolution data to avoid spurious high-frequency noise. End-user needs may also vary depending on whether they are putting climate data into their own impact models, or if they plan to use climate data to inform decisions more broadly.

2.4 References

Adeyeri, OE, P Laux, AE Lawin, and KSA Oyekan. 2020. “Multiple bias-correction of dynamically downscaled CMIP5 climate models temperature projection: a case study of the transboundary Komadugu-Yobe river basin, Lake Chad region, West Africa.” *SN Applied Sciences* 2:1221, https://doi.org/10.1007/s42452-020-3009-4

Akinsanola, AA, C Jung, J Wang, and VR Kotamarthi. 2024. “Evaluation of precipitation across the contiguous United States, Alaska, and Puerto Rico in multi-decadal convection-permitting simulations.” *Scientific Reports* 14(1): 1238, https://doi.org/10.1038/s41598-024-51714-3

Baumberger, C, R Knutti, and G Hirsch Hadorn. 2017. “Building confidence in climate model projections: an analysis of inferences from fit.” *Wiley Interdisciplinary Reviews: Climate Change* 8(3): e454, https://doi.org/10.1002/wcc.454

Bengtsson, L, KI Hodges, and E Roeckner. 2006. “Storm tracks and climate change.” *Journal of Climate* 19(15): 3518–3543, https://doi.org/10.1175/JCL13815.1

Bowden, J., TL Otte, CG Nolte, and MJ Otte. 2012. “Examining interior grid nudging techniques using two-way nesting in the WRF model for regional climate modeling.” *Journal of Climate*
25(8): 2805–2823, https://doi.org/10.1175/JCLI-D-11-00167.1

Chen, X, LR Leung, Y Gao, Y Liu, and M Wigmosta. 2023. “Sharpening of cold-season storms over the western United States.” *Nature Climate Change* 13(2): 167–173, https://doi.org/10.1038/s41558-022-01578-0

CORDEX-SAT. 2020. CORDEX-CMIP5 Simulations. https://cordex.org/data-access/cordex-cmip5-data/cordex-cmip5-simulations/ (Accessed: 15 September 2024).

Donahue, AS, PM Caldwell, L Bertagna, H Beydoun, PA Bogenschutz, A Bradley, TC Clevenger, JG Foucar, JC Golaz, O Guba, WM Hannah, BR Hillman, JN Johnson, N Keen, W Lin, B Singh, S Sreepathi, MA Taylor, J Tian, CR Terai, PA Ullrich, X Yuan, and Y Zhang. 2024. “To exascale and beyond–The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales.” *Journal of Advances in Modeling Earth System*s 16(7): e2024MS004314, https://doi.org/10.1029/2024MS004314

Fandrich, KM, OE Timm, C Zhang, and TW Giambelluca. 2022. “Dynamical Downscaling of Near-Term (2026–2035) Climate Variability and Change for the Main Hawaiian Islands”. *Journal of Geophysical Research – Atmospheres* 127(2): e2021JD035684, https://doi.org/10.1029/2021JD035684

François, B, M Vrac, AJ Cannon, Y Robin, and D Allard. 2020. “Multivariate bias corrections of climate simulations: which benefits for which losses?” *Earth System Dynamics* 11(2): 537–562, https://doi.org/10.5194/esd-11-537-2020

Fox-Rabinovitz, MS, EH Berbery, LL Takacs, and RC Govindaraju. 2005. “A multiyear ensemble simulation of the US climate with a stretched-grid GCM.” *Monthly Weather Review* 133(9): 2505–2525, https://doi.org/10.1175/MWR2956.1

Fu, Y, H Bai, Y Cai, W Yang, and Y Li. 2024. “Optimal configuration method of demand-side flexible resources for enhancing renewable energy integration.” *Scientific Reports* 14(1): 7658, https://doi.org/10.1038/s41598-024-58266-6

Gates, WL. 1985. “The use of general circulation models in the analysis of the ecosystem impacts of climatic change.” *Climatic Change* 7(3): 267–284, https://doi.org/10.1007/BF00144171

Gutmann, ED, JJ Hamman, MP Clark, T Eidhammer, AW Wood, and JR Arnold. 2022. "En-GARD: A statistical downscaling framework to produce and test large ensembles of climate projections." *Journal of Hydrometeorology* 23(10): 1545–1561, https://doi.org/10.1175/JHM-D-21-0142.1

Hall, A., Rahimi, S., Norris, J., Ban, N., Siler, N., Leung, R. L., Ullrich, P., Reed, K., Prein, A. and Qian, Y., 2024. An evaluation of dynamical downscaling methods used to project regional climate change. Journal of Geophysical Research Atmospheres (in press). https://doi.org/10.1002/2016JD024796

Hayhoe, K, A Stoner, DJ Wuebbles, and I Scott-Fleming. 2023. "STAR-ESDM: A Generalizable Approach to Generating High-Resolution Climate Projections through Signal Decomposition." *ESS Open Archive*, https://doi.org/10.22541/essoar.169462036.65393270/v1

Hobeichi, S, N Nishant, Y Shao, G Abramowitz, A Pitman, S Sherwood, C Bishop, and S Green. 2023. "Using machine learning to cut the cost of dynamical downscaling." *Earth's Future* 11(3): e2022EF003291, https://doi.org/10.1029/2022EF003291

Huang, X, DL Swain, and AD Hall. 2020. "Future precipitation increase from very high resolution ensemble downscaling of extreme atmospheric river storms in California." *Science Advances* 6(29): eaba1323, https://doi.org/10.1126/sciadv.aba1323

Huntingford, C, ES Jeffers, MB Bonsall, HM Christensen, T Lees, and H Yang. 2019. "Machine learning and artificial intelligence to aid climate change research and preparedness." *Environmental Research Letters* 14(12): 124007, https://doi.org/10.1088/1748-9326/ab4e55

Jeevanjee, N, and DM Romps. 2018. "Mean precipitation change from a deepening troposphere." *Proceedings of the National Academy of Sciences of the United States of America* 115(45): 11465–11470, https://doi.org/10.1073/pnas.1720683115

Jones, AD, D Rastogi, P Vahmani, AM Stansfield, KA Reed, T Thurber, PA Ullrich, and JS Rice. 2023. "Continental United States climate projections based on thermodynamic modification of historical weather." *Scientific Data* 10(1): 664, https://doi.org/10.1038/s41597-023-02485-5

Komurcu, M, KA Emanuel, M Huber, and RP Acosta. 2018. "High-resolution climate projections for the Northeastern United States using dynamical downscaling at convection-permitting scales." *Earth and Space Science* 5(11): 801–826, https://doi.org/10.1029/2018EA000426

Lafferty, DC, and RL Sriver. 2023. "Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6." *npj Climate and Atmospheric Science* 6(1): 158, https://doi.org/10.1038/s41612-023-00486-0

Ling, F, Z Lu, JJ Luo, SK Bahera, S. D Jin B Pan, H Jiang, and T Yamagata. Diffusion model-based probabilistic downscaling of 190-year East Asian Climate reconstruction. https://arxiv.org/pdf/2402.06646.pdf

Liu, W, PA Ullrich, J Li, C Zarzycki, PM Caldwell, LR Leung, and Y Qian. 2023. "The June 2012 North American derecho: A testbed for evaluating regional and global climate modeling systems at cloud-resolving scales." *Journal of Advances in Modeling Earth Systems* 15(4): e2022MS003358, https://doi.org/10.1029/2022MS003358

Maraun, D. 2013. "Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue." *Journal of Climate* 26(6): 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1

McGinnis, S, and L Mearns. 2021. “Building a climate service for North America based on the NA-CORDEX data archive.” *Climate Services* 22: 100233, https://doi.org/10.1016/j.cliser.2021.100233

Mearns, LO, MS Bukovsky, SC Pryor, and V Magaña. 2014. “Downscaling of Climate Information.” In *Climate Change in North America*, edited by George Ohring, 201–250. Springer International Publishing.

Mizukami et al 2022. “New projections of 21st century climate and hydrology for Alaska and Hawai‘i.” *Climate Services* 27,:100312, https://doi.org/10.1016/j.cliser.2022.100312

Molinari, J, and M Dudek. 1992. “Parameterization of Convective Precipitation in Mesoscale Numerical Models: A Critical Review.” *Monthly Weather Review* 120: 326–344, https://doi.org/10.1175/1520-0493(1992)120<0326:POCPIM>2.0.CO;2

O’Gorman, PA. 2010. “Understanding the varied response of the extratropical storm tracks to climate change.” *Proceedings of the National Academy of Sciences of the United States of America* 107(45): 19176–19180, https://doi.org/10.1073/pnas.1011547107

Otte, TL, CG Nolte, MJ Otte, and JH Bowden. 2012. “Does nudging squelch the extremes in regional climate modeling?” *Journal of Climate* 25(20): 7046–7066, https://doi.org/10.1175/JCLI-D-12-00048.1

Pierce, DW, DR Cayan, and BL Thrasher. 2014. “Statistical downscaling using localized constructed analogs (LOCA).” *Journal of Hydrometeorology* 15(6): 2558–2585, https://doi.org/10.1175/JHM-D-14-0082.1

Pierce, DW, and DR Cayan. 2016. “Downscaling humidity with localized constructed analogs (LOCA) over the conterminous United States.” *Climate Dynamics* 47: 411–431, https://doi.org/10.1007/s00382-915-2845-1

Pierce, DW, DR Cayan, DR Feldman, and MD Risser. 2023. “Future increases in North American Extreme Precipitation in CMIP6 downscaled with LOCA.” *Journal of Hydrometeorology* 24(5): 951–975, https://doi.org/10.1175/JHM-D-22-0194.1

Prein, AF, RM Rasmussen, K Ikeda, C Liu, MP Clark, and GJ Holland. 2017. “The future intensification of hourly precipitation extremes.” *Nature Climate Change* 7(1): 48–52, https://doi.org/10.1038/nclimate3168

Rahimi, S, L Huang, J Norris, A Hall, N Goldenson, W Krantz, B Bass, C Thackeray, H Lin, D Chen, E Dennis, E Collins, ZJ Lebo, E Slinskey, S Graves, S Biyani, B Wang, S Cropper, and the UCLA Center for Climate Science Team. 2024a. “An overview of the Western United States Dynamically Downscaled Dataset (WUS-D3).” *Geoscientific Model Development* 17(6): 2265–2286, https://doi.org/10.5194/gmd-17-2265-2024

Rahimi, S, L Huang, J Norris, A Hall, N Goldenson, M Risser, DR Feldman, ZJ Lebo, E Dennis, and C Thackeray. 2024b. “Understanding the cascade: Removing GCM biases improves dynamically downscaled climate projections.” *Geophysical Research Letters* 51(9): e2023GL106264, https://doi.org/10.1029/2023GL106264

Rasmussen, R, K Ikeda, C Liu, D Gochis, M Clark, A Dai, E Gutmann, J Dudhia, F Chen, M Barlage, and D Yates. 2014. “Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations.” *Journal of Hydrometeorology* 15(3): 1091–1116, https://doi.org/10.1175/JHM-D-13-0118.1

Rio, C, AD Del Genio, and F Hourdin. 2019. “Ongoing Breakthroughs in Convective Parameterization.” *Current Climate Change Reports* 5: 95–111. https://doi.org/10.1007/s40641-019-00127-w

Risser, M.D., Rahimi, S., Goldenson, N., Hall, A., Lebo, Z.J. and Feldman, D.R., 2024. Is Bias Correction in Dynamical Downscaling Defensible?. *Geophysical Research Letters*, *51*(10), p.e2023GL105979. https://doi.org/10.1029/2023GL105979

Stengel, K, A Glaws, D Hettinger, and RN King. 2020. “Adversarial super-resolution of climatological wind and solar data.” *Proceedings of the National Academy of Sciences of the United States of America* 117(29): 16805–16815, https://doi.org/10.1073/pnas.1918964117

Tang, Q, SA Klein, S Xie, W Lin, JC Golaz, EL Roesler, MA Taylor, PJ Rasch, DC Bader, LK Berg, and P Caldwell. 2019. “Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling.” *Geoscientific Model Development* 12(7): 2679–2706, https://doi.org/10.5194/gmd-12-2679-2019

Thrasher, B, W Wang, A Michaelis, F Melton, T Lee, and R Nemani. 2022. “NASA global daily downscaled projections CMIP6.” *Scientific Data* 9(1): 262, https://doi.org/10.1038/s41597-022-01393-4

Ulbrich, U, JG Pinto, H Kupfer, GC Leckebusch, T Spangehl, and M Reyers. 2008. “Changing Northern Hemisphere storm tracks in an ensemble of IPCC climate change simulations.” *Journal of Climate* 21(8): 1669–1679, https://doi.org/10.1175/2007JCLI1992.1

Ullrich, PA. 2023. Validation of LOCA2 and STAR-ESDM Statistically Downscaled Products. LLNL-TR-856094: 1085545, https://doi.org/10.2172/2202926

Wang, J, and VR Kotamarthi. 2015. “High-resolution dynamically downscaled projections of precipitation in the mid and late 21st century over North America.” *Earth's Future* 3(7) 268–288, https://doi.org/10.1002/2015EF000304

Wang, J, Z Liu, I Foster, W Chang, R Kettimuthu, and VR Kotamarthi. 2021. “Fast and accurate learned multiresolution dynamical downscaling for precipitation.” *Geoscientific Model Development* 14(10): 6355–6372, https://doi.org/10.5194/gmd-14-6355-2021

Wilby, RL, H Hassan, and K Hanaki. 1998. “Statistical downscaling of hydrometeorological variables using general circulation model output.” *Journal of Hydrology* 205(1-2): 1–19, https://doi.org/10.1016/S0022-1694(97)00130-3

Wilcke, RAI, T Mendlik, and A Gobiet. 2013. “Multi-variable error correction of regional climate models.” *Climatic Change* 120: 871–887, https://doi.org/10.1007/s10584-013-0845-x

Wilkinson, M, M Dumontier, I Aalbersberg, G Appleton, M Axton, A Baak, N Blomberg, J-W Boiten, L Bonino da Silva Santos, PE Bourne, J Bouwman, AJ Brookes, T Clark, M Crosas, I Dillo, O Dumon, S Edmunds, CT Evelo, R Finkers, A Gonzalez-Beltran, AJG Gray, P Groth, C Goble, JS Grethe, J Heringa, PAC ’t Hoen, R Hooft, T Kuhn, R Kok, Joost Kok, SJ Lusher, ME Martone, A Mons, AL Packer, B Persson, P Rocca-Serra, M Roos, R van Schaik, S-A Sansone, E Schultes, T Sengstag, T Slater, G Strawn, MA Swertz, M Thompson, J van der Lei, E van Mulligen, J Velterop, A Waagmeester, P Wittenburg, K Wolstencroft, J Zhao, and B Mons. 2016. “The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific Data* 3: 160018, https://doi.org/10.1038/sdata.2016.18

Wood, AW, EP Maurer, A Kumar, and DP Lettenmaier. 2002. “Long-range experimental hydrologic forecasting for the eastern United States.” *Journal of Geophysical Research – Atmospheres* 107(D20): ACL-6-1–ACL-6-15, https://doi.org/10.1029/2001JD000659

Xu, Z, and Z-L Yang. 2015. “A new dynamical downscaling approach with GCM bias corrections and spectral nudging.” *Journal of Geophysical Research – Atmospheres* 120(8): 3063–3084, https://doi.org/10.1002/2014JD022958

Yang, X, EF Wood, J Sheffield, L Ren, M Zhang, and Y Wang. 2018. “Bias correction of historical and future simulations of precipitation and temperature for China from CMIP5 models.” *Journal of Hydrometeorology* 19(3): 609–623, https://doi.org/10.1175/JHM-D-17-0180.1

Yin, JH. 2005. “A consistent poleward shift of the storm tracks in simulations of 21st century climate.” *Geophysical Research Letters* 32(18): L18701, https://doi.org/10.1029/2005GL023684

Zarzycki, CM, C Jablonowski, DR Thatcher, and MA Taylor. 2015. “Effects of localized grid refinement on the general circulation and climatology in the Community Atmosphere Model.” *Journal of Climate* 28(7): 2777–2803, https://doi.org/10.1175/JCLI-D-14-00599.1

3.0 The State of Knowledge on Users' Climate Information Needs

Lead Authors:

Kripa Jagannathan, Lawrence Berkeley National Laboratory, kajagannathan@lbl.gov

Andy Jones, Lawrence Berkeley National Laboratory, adjones@lbl.gov

Contributing Authors:

Jen Bewley, Institute for Defense Analysis, jbewley@ida.org

Art Degaetano, Cornell University, atd2@cornell.edu

Ankur Dixit, Cornell University, ad933@cornell.edu

Keith Dixon, National Oceanic and Atmospheric Administration, keith.dixon@noaa.gov

Jeremy Giovando, U.S. Army Corps of Engineers, Jeremy.J.Giovando@usace.army.mil

Geneva Gray, U.S. Environmental Protection Agency, Gray.Geneva@epa.gov

Ethan Gutmann, National Center for Atmospheric Research, gutmann@ucar.edu

Rao Kotamarthi, Argonne National Laboratory, vrkotamarthi@anl.gov

Hugo Lee, NASA Jet Propulsion Laboratory, huikyo.lee@jpl.nasa.gov

Jeremy Littell, U.S. Geological Survey, jlittell@usgs.gov

Andrew Schwarz, California Department of Water Resources, andrew.schwarz@water.ca.gov

Oliver Ellison Timm, University of Albany, oelisontimm@albany.edu

Nathalie Voisin, Pacific Northwest National Laboratory, nathalie.voisin@pnnl.gov

Adrienne Wootten, University of Oklahoma, amwootte@ou.edu

3.1 Chapter Summary

In spite of rapid climate science advances, most climate data users continue to find it extremely difficult to appropriately examine and work with climate data. Co-production and other such collaborative approaches that bring together data producers and different kinds of users can support better understanding of decision-making contexts and improve the actionability of climate data. However, co-production is resource and time intensive because it needs to be expert-facilitated and collaborative across many disciplines.

3.2 Background

This chapter discusses what is known or not known about various stakeholder groups' needs for climate information. The sub-sections discuss types of climate information users, types of actionable climate information, types of use-cases or decisions for which climate information is needed and finally federal agencies that are actively working towards better identifying and understanding different stakeholder groups' needs for climate information.

3.3 Types of Climate Data Users

Many varied groups use climate information and data products, but tend to have very different use-cases (or needs) as well as varied technical capabilities. This makes it difficult to understand which climate data products are the best fit in these different contexts (Bessembinder et al. 2019). Groups who use climate information can range from researchers through planners and decision-makers to tribal entities and public consumers of data. Figure 6 showcases one such categorization of different data user groups that was developed for the California Energy Commission funded Cal-Adapt Analytics Engine project.

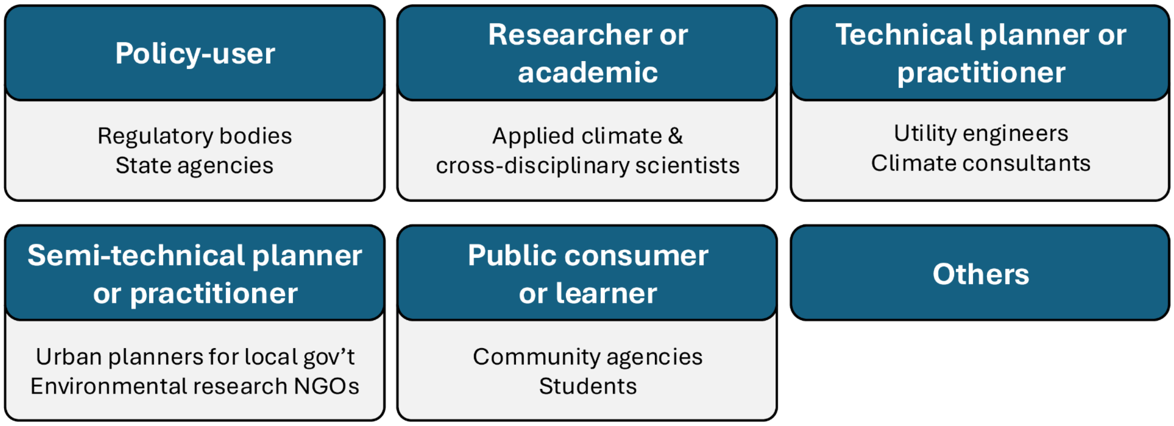


Figure 6. Categorization of different climate data user groups developed for the Cal-Adapt Analytics Engine project and select examples. Figure and categorization developed by: Justine Bui, Grace Di Cecco, Ashley Conrad-Saydah, Nancy Freitas, Kripa Jagannathan, Nancy Thomas, Owen Doherty, Mark Koenig, and the Cal-Adapt Analytics Engine Team (based on preliminary results from ongoing work).

Technical proficiency in using climate data varies both within and among these user groups (Bessembinder et al. 2019, Raaphorst et al. 2020). Particularly, the diversity in needs and capabilities among different planning and decision-making communities can be quite large, as their needs depend on the type of decision as well as each group's mandates, missions, and risk framing. These user groups can range from relatively decentralized community/municipal planning to deliberate, structured decision-making in federal agencies managing public trust resources. Some have (or can have) access to technical personnel and resources who can seamlessly incorporate output from climate models or downscaled projections into their own modeling (e.g., consultative relationships are fairly common in water resources). Others have severely limited technical capacity and rely heavily on ongoing partnerships to obtain, characterize, and use data and derived information to make decisions. Furthermore, the needs and capabilities of these user groups evolve over time.

While decision contexts for each individual agency or group might be unique, many workshop attendees discussed the need for a better understanding and categorization of the types of users and their needs based on broad data demand categories that branch into more decision contexts. The need to move away from the data producer versus data user dichotomic categorization was also discussed. Specifically, users can also be engaged partners in data production processes, particularly with the increase of collaborative scientific processes such as co-production (Bremer and Meisch 2017, Lemos et al. 2012).

3.4 Types of Actionable Information

The term "actionable" is often not defined by the type of information, but rather primarily by the decision context, including governance and decision-making processes in which the information will be used. The same type of information may be actionable in one context, but not in another. However, some common information needs exist and have been documented in various types of literature (Hackenbruch et al. 2017, Vincent et al. 2020, Jagannathan et al. 2022). Figure 7 summarizes one such typology of actionable climate information. Almost universally, users need repeatable, accessible, thoroughly vetted, and defensible projections of climate change. Many users almost invariably need information on (a) understanding how averages and common event likelihoods are shifting and the relative probability/timing of those shifts, (b) extremes (time/space events that were historically rare or are novel in the future, and/or have the biggest economic or ecosystem impacts) and (c) downscaled climate data not just for temperature and precipitation, but a host of related, derived physical and hydrologic variables (or decision-relevant metrics) that actually drive Earth system processes and impact actual management endpoints, such as droughts, floods, runoff, streamflow, soil moisture, snow, and permafrost. Researchers have identified what these metrics and variables may look like in different contexts (for instance in Vincent et al. 2020, Jagannathan et al. 2021, Reed et al. 2022).

Actionable information for users needs to be specific to the domain and at a scale/resolution that represents decision-makers' purview/decision context(s). The data need to cover a range of scenarios and models that capture the main sources of projection uncertainty to arrive at a plausible range of futures that match the risk framing required by the decision-maker and their mandate(s). Workshop discussions also noted that a "plausible range" of futures cannot just be scientifically determined; rather, this range needs to be examined in a risk management context, with an assessment of relative likelihood/probability/quasi-probability/level of concern for different points on the range. Overall, there is a need to iterate between the decision-maker's need for true probabilistic risk assessment and scientific limitations on providing such probability assessments.

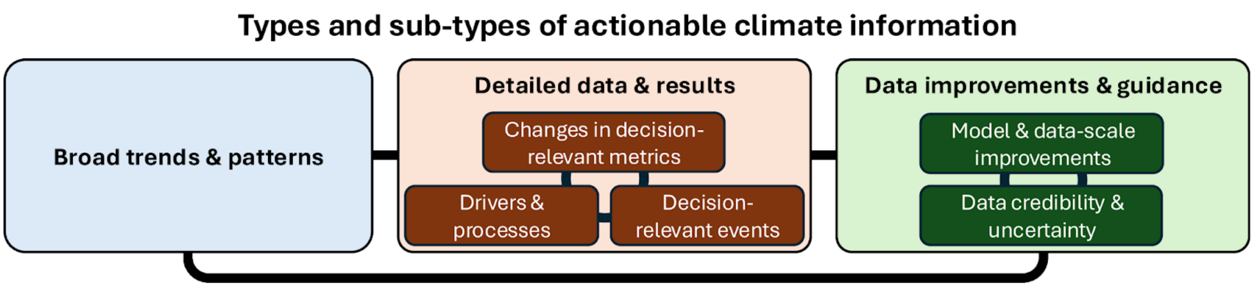


Figure 7. Typology of actionable climate information derived from the iterative co-production engagements conducted in the DOE-funded HyperFACETS project. Based on Jagannathan et al. 2022.

3.5 Types of Uses of Climate Information

Often users are requested by data providers to elucidate the types of uses they intend for the climate data, i.e., the decision/s that they intend to make with the climate information. However, the use and decision landscape is extremely vast and complicated and often difficult to quickly summarize. For instance, the types of use or decisions vary by: sector, the type of management issue within the sector, and the numerous individual decisions or use-cases within each of these sector-specific management issues (see Figure 8 for an example). Furthermore, each decision is not a discrete event but a dynamic and long-drawn process requiring different types of information at different times. Figure 9 illustrates one typology of uses of climate information developed through an analysis of co-production engagements between climate scientists and water, energy, and land managers across the U.S.

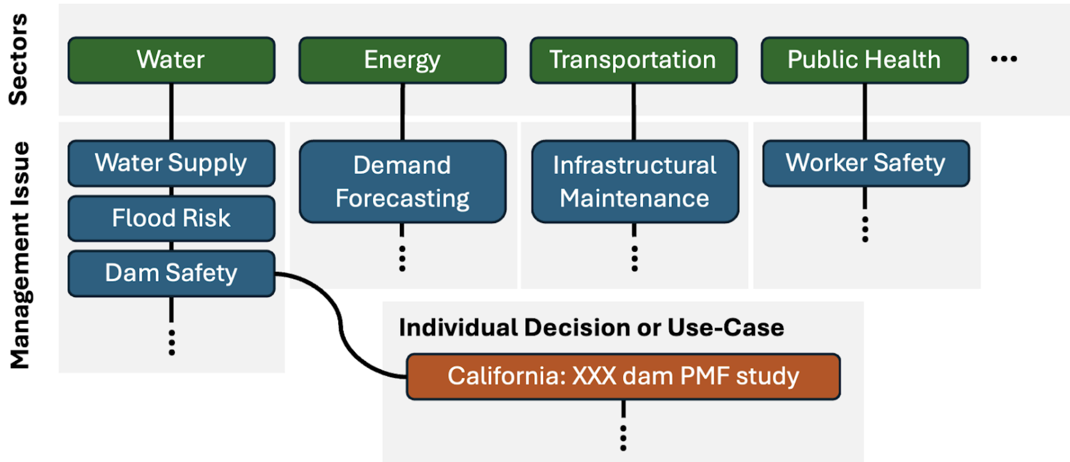


Figure 8. **Potential types of use or decisions by sector, management issue, and individual decision or use-case.**

Figure 9. **Typology of uses of climate information derived from the iterative co-production engagements conducted in the DOE-funded HyperFACETS project. Reproduced from Jagannathan et al. 2022.**

3.6 Boundary Agencies Engaging with Climate Information Users

Making information actionable often requires more than providing the right climate data and projections. It needs investment in relationships and resources for deliberate and iterative boundary spanning between researchers/purveyors and decision-makers/users. Boundary agencies, who aim to connect and bridge the boundaries between information providers and users, are extremely important players in this space. This boundary spanning includes undertaking nuanced needs assessments for the different user communities, and facilitating partnerships where data are co-developed, translated, and interpreted for different use contexts. Several specialized agencies have been tasked with undertaking such boundary-spanning activities.

Some prominent national-level agencies doing this work include the NOAA Climate Adaptation Programs (NOAA-CAPs, formerly called the NOAA RISAs), who have been doing this since the mid-1990s and represent a compelling set of regional partnerships that have substantially contributed to knowledge of stakeholder/rightsholder elicitation and needs assessments. NOAA's National Centers for Environmental Information (NCEI) also has a set of regional Climate Service Directors engaging with end-users to ascertain climate information needs. The U.S. Geological Survey (USGS) Climate Adaptation Science Centers (CASCs) engage with end-users whose missions/mandates include adaptation to climate change within the U.S. Department of the Interior and partners. The CASCs also engage with the Bureau of Indian Affairs and tribal entities seeking climate science as input into adaptation based on multiple sources of knowledge. The U.S. Department of Agriculture's (USDA) regional Climate Hubs are another set of prominent boundary agencies, working across the USDA and with partners to support climate-informed decisions for robust agriculture, healthy forests, and resilient communities. In addition to these federal agencies, several large climate research projects such as the DOE-funded Urban Integrated Field Labs (UIFL) and HyperFACETS, as well as state-level data-sharing and curation platforms such as the Cal-Adapt and Cal-Adapt Analytics Engine, are also prominent in understanding user needs and providing actionable climate information and tools. Increasingly, private climate service providers and climate consulting firms are also playing a significant role in this space. Collaborations with these boundary agencies, as well as reviewing resources developed by these agencies such as user guides or needs assessment reports, can be extremely valuable for the data producer community; for example, the climate data user guide from the Electric Power Research Institute (EPRI 2024), the consumer report for climate information from Great Lakes Integrated Sciences and Assessments (GLISA; Briley et al. 2020), and the White House Office of Science and Technology Policy (OSTP) guide on selecting climate information (OSTP 2023).

3.7 Progress in Meeting User Needs for Climate Information

3.7.1 Collaborative Approaches to Improve the Actionability of Climate Data

Collaborative data production processes that iteratively engage with potential users, and incorporate users' experiences and knowledge into the process, have been shown to increase the actionability of data for decision-making. Over the last 20 years, more nuanced characterizations of these approaches have emerged, and have increased our ability to work with a wide range of users in varied engagement modes such as through consultative partnerships, meaningful collaboration, or iterative co-production (see Meadow et al. 2015, Bremer and Meisch 2017). There is increasing evidence across different contexts that collaborative processes tend to improve the credibility, legitimacy, and salience of climate data, and hence improves the use of climate projections in decision-making.

There is also an increased recognition that well-designed-and-executed collaborative processes require dedicated time, capacity, and resources, and also need a specific set of expertise, skills, and capacities (Lemos et al. 2014). More resources are needed to increase and better support these critical boundary-spanning activities (Goodrich et al. 2020). Therefore, emerging work is also focusing on how to scale up such co-production or collaborative approaches in a cost- and resource-effective manner. Although every decision context is unique, there is also an acknowledgement that it is not possible to intensively engage with every potential user and decision context. Many scholars are conducting "meta" studies of several co-production efforts and starting to develop generalizable frameworks (and actionable recommendations) of the broad types of data products, decision contexts, engagement approaches, and institutional contexts that can be applicable to multiple users and decision contexts (Bamzai-Dodson et al. 2021).

3.7.2 Examples of Collaborative Efforts

Many promising and successful examples of producer-user collaborations in different agencies and regional contexts have recently emerged and were highlighted at the workshop. It was noted that collaborative efforts can be at the project, programmatic, or institutional level (such as in the NOAA-CAPs, or USGS CASCS, where the entire institutional structures are also collaborative and specifically intended to develop actionable climate information partnerships). Projects ranged from development of nation-wide or state-wide climate data tools and portals to individual projects working with a group/groups of users to develop actionable information. Some projects that were discussed include the DOE-funded HyperFACETS, IFLs and Climate Risk and Resilience Portal; South Central CASC's Edwards Aquifer Authority project; Alaska CASCs programs on provision of climate information/services; Weather Effects on the Lifecycle of DOD Equipment Replacement or WELDER project; and the Cal-Adapt Analytics Engine.

3.8 Remaining Gaps in Meeting User Needs

Four categories of gaps were identified as needing further work.

3.8.1 Gaps in Available Data and Tools

The workshop attendees noted that despite several advances in the provision of climate data, some gaps in regional data (e.g., data and methods for Alaska and the Pacific islands) as well as decision-relevant variables and resolutions, still persisted. In addition, there are limited user-friendly tools and analytics to help users parse through and work with the incredibly large amounts of downscaled climate data available. Other chapters of this report further elaborate on these data and tool gaps.

3.8.2 Lack of Guidance on Appropriate Use of Climate Data for Decision Applications

Even the most sophisticated of climate data users in the practitioner community often find it extremely difficult to navigate the complexities of the diverse types of climate data products available to them. Most users report that there are no transparent and user-centric guidances (i.e., dos and don'ts) on appropriate use of climate data for decision applications (Vano et al. 2018). Most recommendations on climate data use tend to be extremely academic without much understanding of the applicability within the bounds and limitations of a decision-making space. Yet, users are often called out for using data inappropriately. The workshop attendees noted the lack of actionable guidance as one of the biggest gaps in the use of climate information, and one of the biggest research and scientific gaps in climate science today.

Users need better guidance, first on choice of data products (i.e., when to use each type of data), and on the pros and cons of using different data products for different applications. They also need guidance on the most 'scientifically appropriate but also practically operational' way of using climate information for specific decisions. Guidance on the following is often lacking:

- Depth of analysis needed for different application types (e.g., when can you rely on an online ranking or visualization tools that provide broad regional summaries or statistics or vulnerability rankings, versus when there is a need for deeper analysis of data).
- Best practices and well-characterized approaches in integrating climate model outputs into impacts and management modeling that transparently identify the risks and the additional uncertainties during such coupling.
- Choices of resolution or downscaling approaches:
  - Pros and cons of choosing between different downscaling datasets
  - When finer-scale resolution is required or necessary
  - Comparison of uncertainties between various approaches
  - Needs for physical consistency among variables.
- Approaches to characterize and understand different sources and magnitudes of uncertainties and how they translate to choices of scenarios, models, etc. This can include use of various risk-relevant metrics such as likelihood, probability density function (PDF), and confidence level and prediction intervals, combined with scientific metrics of uncertainties.

Such guidance needs to be developed collaboratively between producers, users, regulatory agencies, and boundary agencies. The role of boundary agencies and boundary chains in helping to translate these guidelines accurately into different decision contexts is critical. Local climate boundary agencies might need to be empowered and strengthened so that they can serve as a valuable conduit between the producer and user communities. Some workshop participants also brought up the relatively recent proliferation of using AI for climate services, and the need for guidance on this space.

3.8.3 Lack of Credibility Evaluations from a Practitioner's Standpoint

Another gap for users that came up prominently during the workshop was that credibility evaluations of climate data are often conducted purely from a climate science perspective without consideration of credibility for use in decisions. It was brought up that credibility may mean different things to different producers, users, and data evaluators; hence credibility could also be context dependent. As one workshop participant mentioned: "*10TB of highly resolved projection that is highly publishable may be useless in the wrong decision context, so definitions of credibility can vary*." Among the academic community, there are opportunities to develop evaluation methods to better characterize the skill and applicability of datasets for various use cases.

More specifically, it was discussed that, in the interest of scientific confirmation and diversity of available products, in many instances there should not be a single designated 'official' set of downscaled projections for any and all purposes. At the same time, in a tangled ecosystem of semi-authoritative products, it is difficult to say who decides the standards for data being fit-for-purpose, and there likely should be a structured, facilitated, collaborative effort to address this issue among various stakeholders. Some specific aspects of credibility assessments that came up included the need for a standardized evaluation workflow (as discussed in chapter 4), a standard set of questions to ask to understand how a stakeholder defines credibility, and a standardization of metadata information that all data products must have: i.e., the temporal and spatial resolution, map projection, time periods, quality control information, narrative about the data set, and so on that should be included with data set documentation to assist users, translators, and evaluators with assessing and using projection data (as discussed in chapter 2).

3.8.4 Gaps in User Engagement and Resources to Support it

Despite progress in user engagement, discussions nonetheless identified many remaining gaps. For instance, it was brought up that most data producers' understanding of actual decision contexts in which the data may be used was very limited. Therefore there is still a need for data producers to better engage with and understand the different users' workflow and contexts. Further, it was brought up that not all engagements are successful, and there is only emerging understanding on what types of engagements work and which ones do not work. There is a need for more evaluation/synthesis of engagement approaches and resulting/implied best practices. There is emerging evidence that successful user engagement often requires skilled expert facilitators who have competence both in climate data/developer representation and the decision context/stakeholder/user representation; however, such expertise is hard to come-by and training to build such competencies is very limited. Further, climate data user-mapping (i.e., mapping of stakeholder questions to particular datasets) can be a complicated endeavor, and it is often hard to even identify who in an organization actually uses the climate data and makes decisions on data choice. Lastly, the conflict of interest that might arise when climate data provision services for-profit activities was also brought up. While private data providers will have an important role in enabling accessibility of the data, there may also be cases where there are some potential conflicts of interest.

3.9 Way Forward

Given the uneveness in available computational, financial, and human resources across sectors, utilities, agencies, and companies, more research and evaluation of how climate data are actually being used, and who is using it, is necessary. New climate data sets need to be informed by the actual needs and questions from the practitioners, rather than from aspirations on how they could be used. Furthermore, the role of ML chatbots or other ChatGPT-like ML tools in the provision of data also needs to be explored (and critiqued). AI and machine learning approaches can potentially be powerful tools to assist in meeting end-user needs, but there is a need for careful consideration/development of ethical standards and guardrails for AI use.

Improving the provision of useful and usable climate data for a variety of research, decision-making, and other applications requires a collaborative relationship between data producers, evaluators, users, and boundary agencies that work to connect data providers and users. One potential approach is a community of practice (COP), but if a new COP is being developed, then it needs to have a clear and well-defined goal, and adequate resources to help sustain and maintain it. Best practices and literature on effectively designing and sustaining such COPs should also be followed (e.g., Miles et al. 2006, Page and Dilling 2019). Connections (and overlap) with other communities of practice such as the Science for Climate Action Network (SCAN), the Cal-Adapt Analytics Engine, DOE's Multi-Sector Dynamics research-to-operations-to-research community of practice (MSD R2O2R), the Electric Power Research Institute's ClimateREADi group, the Integrated Hydro-Terrestrial Modelling group (IHTM), and the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI), should be explored.

3.10 References

Bamzai-Dodson, A, AE Cravens, AA Wade, and RA McPherson. 2021. "Engaging with stakeholders to produce actionable science: a framework and guidance." *Weather, Climate, and Society* 13(4): 1027–1041, https://doi.org/10.1175/WCAS-D-21-0046.1

Bessembinder, J, M Terrado, C Hewitt, N Garrett, L Kotova, M Buonocore, and R Groenland. 2019. "Need for a common typology of climate services." *Climate Services* 16: 100135, https://doi.org/10.1016/j.cliser.2019.100135

Bremer, S, and S Meisch. 2017. "Co-production in climate change research: reviewing different perspectives." *Wiley Interdisciplinary Reviews: Climate Change* 8(6): e482, https://doi.org/10.1002/wcc.482

Briley, L, R Kelly, ED Blackmer, AV Troncoso, RB Rood, J Andresen, and MC Lemos. 2020. "Increasing the usability of climate models through the use of consumer-report-style resources for decision-making." *Bulletin of the American Meteorological Society* 101(10): E1709–E1717, https://doi.org/10.1175/BAMS-D-19-0099.1

EPRI. 2024. Climate Data Users Guide. V1.0.0. 3002028078. Palo Alto, California. https://www.epri.com/research/products/000000003002028078

Goodrich, KA, KD Sjostrom, C Vaughan, L Nichols, A Bednarek, and MC Lemos. 2020. "Who are boundary spanners and how can we support them in making knowledge more actionable in sustainability fields?" *Current Opinion in Environmental Sustainability* 42: 45–51, https://doi.org/10.1016/j.cosust.2020.01.001

Hackenbruch, J, T Kunz-Plapp, S Müller, and JW Schipper. 2017. "Tailoring climate parameters to information needs for local adaptation to climate change". *Climate* 5(2): 25, https://doi.org/10.3390/cli5020025

Jagannathan, K, S Buddhavarapu, PA Ullrich, AD Jones, and HyperFACETS Project Team. 2023. "Typologies of actionable climate information and its use." *Global Environmental Change* 82: 102732, https://doi.org/10.1016/j.gloenvcha.2023.102732

Jagannathan, K, AD Jones, and I Ray. 2021. "The making of a metric: Co-producing decision-relevant climate science." *Bulletin of the American Meteorological Society* 102(8): E1579–E1590, https://doi.org/10.1175/BAMS-D-19-0296.1

Lemos, MC, CJ Kirchhoff, and V Ramprasad. 2012. "Narrowing the climate information usability gap." *Nature Climate Change* 2(11): 789–794, https://doi.org/10.1038/nclimate1614

Lemos, MC, CJ Kirchhoff, SE Kalafatis, D Scavia, and RB Rood. 2014. "Moving climate information off the shelf: boundary chains and the role of RISAs as adaptive organizations." *Weather, Climate, and Society* 6(2): 273–285, https://doi.org/10.1175/WCAS-D-13-00044.1

Meadow, AM, DB Ferguson, Z Guido, A Horangic, G Owen, and T Wall. 2015. "Moving toward the deliberate coproduction of climate science knowledge." *Weather, Climate, and Society* 7(2): 179–191, https://doi.org/10.1175/WCAS-D-14-00050.1

Miles, EL, AK Snover, L Whitely Binder, E Sarachik, P Mote, and N Mantua. 2006. "An approach to designing a national climate service." *Proceedings of the National Academy of Sciences of the United States of America* 103(52): 19616–19623, https://doi.org/10.1073/pnas.0609090103

OSTP 2023. Selecting Climate Information to Use in Climate Risk and Impact Assessments: Guide for Federal Agency Climate Adaptation Planners. White House Office of Science and Technology Policy. Washington, D.C., USA, https://www.whitehouse.gov/wp-content/uploads/2023/03/Guide-on-Selecting-Climate-Information-to-Use-in-Climate-Risk-and-Impact-Assessments.pdf

Page, R, and L Dilling. 2019. "The critical role of communities of practice and peer learning in scaling hydroclimatic information adoption." *Weather, Climate, and Society* 11(4): 851–862, https://doi.org/10.1175/WCAS-D-18-0130.1

Raaphorst, K, G Koers, GJ Ellen, A Oen, B Kalsnes, L van Well, J Koerth, and R van der Brugge. 2020. "Mind the gap: towards a typology of climate service usability gaps." *Sustainability* 12(4): 1512, https://doi.org/10.3390/su12041512

Reed, KA, N Goldenson, R Grotjahn, WJ Gutowski, K Jagannathan, AD Jones, LR Leung, SA McGinnis, SC Pryor, AK Srivastava, and PA Ullrich. 2022. "Metrics as tools for bridging climate science and applications." *Wiley Interdisciplinary Reviews: Climate Change* 13(6): e799, https://doi.org/10.1002/wcc.799

Vano, JA, JR Arnold, B Nijssen, MP Clark, AW Wood, ED Gutmann, N Addor, J Hamman, and F Lehner. 2018. "DOs and DON'Ts for using climate change information for water resource planning and management: guidelines for study design." *Climate Services* 12: 1–13, https://doi.org/10.1016/j.cliser.2018.07.002

Vincent, K, E Archer, R Henriksson, J Pardoe, and N Mittal. 2020. "Reflections on a key component of co-producing climate services: defining climate metrics from user needs." *Climate Services* 20: 100204, https://doi.org/10.1016/j.cliser.2020.100204

4.0 The State of Benchmarking and Evaluation for Regional Climate Data Sets and Projections

Lead Authors:

Hugo Lee, NASA Jet Propulsion Laboratory, huikyo.lee@jpl.nasa.gov

Paul Ullrich, Lawrence Livermore National Laboratory, ullrich4@llnl.gov

Contributing Authors:

Moetasim Ashfaq, Oak Ridge National Laboratory, mashfaq@ornl.gov

Ankur Dixit, Cornell University, ad933@cornell.edu

Keith Dixon, National Oceanic and Atmospheric Administration, keith.dixon@noaa.gov

Ethan Gutmann, National Center for Atmospheric Research, gutmann@ucar.edu

Alex Hall, University of California – Los Angeles, alexhall@g.ucla.edu

Sepideh Khajehei, NASA Ames Research Center, sepideh.khajehei@nasa.gov

David Lafferty, University of Illinois – Urbana-Champaign, davidcl2@illinois.edu

Flavio Lehner, Cornell University, flavio.lehner@cornell.edu

Ruby Leung, Pacific Northwest National Laboratory, ruby.leung@pnnl.gov

Elias Massoud, Oak Ridge National Laboratory, massoudec@ornl.gov

Tanya Spero, U.S. Environmental Protection Agency, spero.tanya@epa.gov

Adrienne M. Wootten, University of Oklahoma, amwootte@ou.edu

4.1 Chapter Summary

Benchmarking and evaluating downscaled data sets against observations and/or with controlled experiments are critical steps that support and enable the use of those data sets in a wide range of applications for resilience and adaptation planning. The global modeling community has established mature and standard approaches to evaluation and benchmarking, but downscaled data producers, translators, and practitioners have not. With new data production efforts underway and novel production techniques emerging, standardized evaluation techniques that are independent of production method are increasingly needed. The components necessary to achieve robust standards for benchmarking and evaluation of downscaled climate projections, as well as some of the research needed to achieve that goal, are presented and discussed.

4.2 Background

This chapter compiles the discussions from the workshop and provides a summary of the current landscape of benchmarking and evaluation for downscaled products. "Benchmarking" refers to the methodology of comparing downscaled products with reference observations (e.g., in situ weather data), observational products (e.g., gridded temperatures and precipitation), or optimal methodologies acknowledged by the climate scientific community. In contrast, "evaluation" denotes a comprehensive analysis focused on determining the reliability, precision, and relevance of downscaled products, while using various metrics. Benchmarking and evaluation are crucial for establishing the ability of climate models to produce useful results for decision-making, both individually and collectively. For example, Figure 10 depicts a common workflow for deriving meaningful conclusions from decision-relevant climate data, where information from benchmarking and evaluation provides an essential early step that shapes the conclusions drawn from those data. Evaluation provides data producers with insights that can guide future model improvements and help users to understand the strengths and limitations of data sets for specific applications. This practice advances the understanding of climate change, where metrics serve as indicators for possible issues within the process representations that can be directly or indirectly relevant to a model's ability to represent a particular impact. When a model exhibits a low metric score or sets of scores, it indicates a need for further qualitative and physical investigations. Evaluation is also used for deriving model weights in ensemble runs, for linking confidence in model results to its representation of historical climate, and for deriving best estimates and uncertainty in future change.

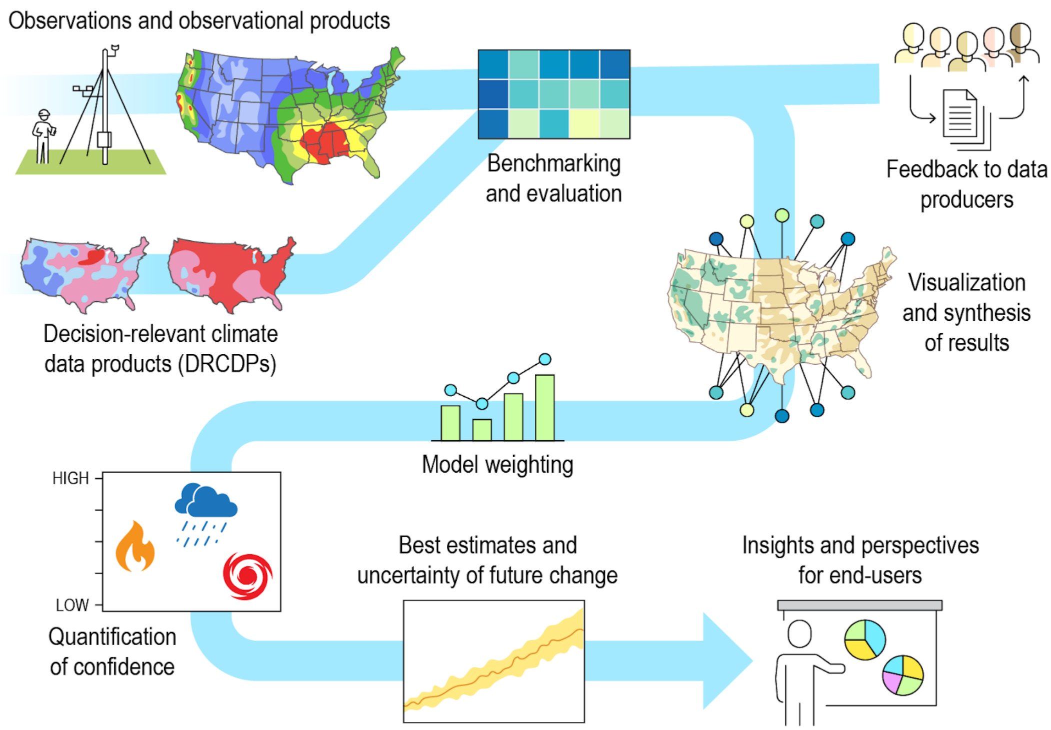


Figure 10. A schematic representation of the evaluation workflow connecting data products to output supporting data producers and end-users.

4.3 Standardized Evaluation of Decision-Relevant Climate Data

Standardized model evaluation for GCMs today is a far more mature discipline, largely due to efforts by DOE's PCMDI in the early 1990s that laid the groundwork for CMIP (Potter et al. 2011). At present, extensive and readily available benchmarking tools (for instance, the PCMDI Metrics Package (Lee et al. 2023) and ESMValTool (Righi et al. 2020)), are frequently used by operational centers for diagnosing model biases and tracking model improvements over time. For example, Figure 11 shows a "portrait plot" for GCMs, in this case for quantifying relative model performance on a variety of climatological metrics. Efforts focused on standardized evaluation have greatly contributed to the tracking of and improvements in GCM performance across CMIP generations as well as building confidence in model projections (Bock et al., 2020).

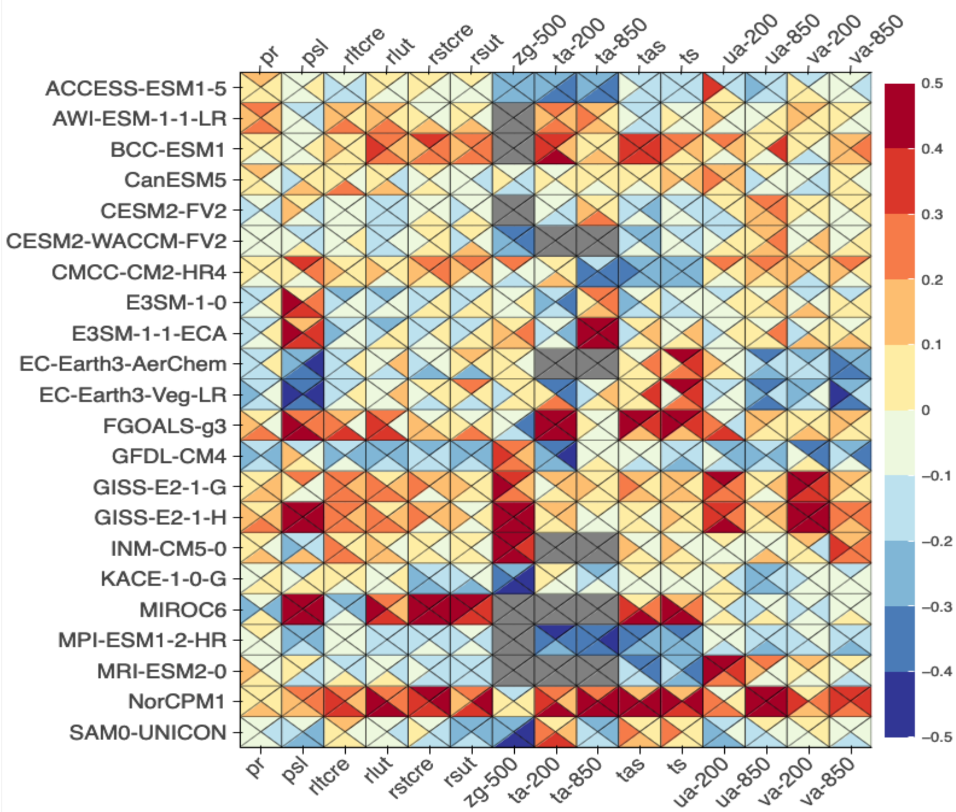


Figure 11. A visualization of GCM performance across standard climatological metrics using a portrait plot. Visualizations such as these are commonplace in GCM evaluation, and could be adapted for use by the regional climate data community.

However, there are notable differences for standardized evaluation of DRCDPs that hinder an exact translation of accumulated GCM diagnostic techniques and experiences to the DRCDP space. First, GCM evaluation primarily addresses evaluation of the simulated climate system, while decision-relevant data needs to be evaluated with end-user needs in mind. Currently, there is no approach to mapping end-user needs to bespoke evaluation approaches and metrics. Second, GCM evaluation is more easily coordinated because the geographic coverage of each model and experiment is the same; the choice of domain extent can be very impactful on DRCDP results and is an integral part of the evaluation process. Finally, the evaluation of DRCDPs must consider the sensitivity of those projections to initial and boundary conditions and/or training data choices.

Despite the features that are unique to DRCDP evaluation, there was general consensus at the workshop that there is a need for the development of a standardized DRCDP evaluation framework for ascertaining the performance of various downscaled products, tracking that performance over time, and allowing data users to readily intercompare and understand differences among products using consistent evaluation criteria with rational, justifiable bases.

Such a framework would consist of a suite of standard metrics and diagnostics, along with prescribed methods for performing that computation that are relatively flexible across methodological choices in the data production process (e.g., downscaling method, choice of grid). This further necessitates a clear articulation and justification of the standard metrics and diagnostics used to perform this evaluation, along with what is learned from such evaluation. Of course, such a standardized evaluation workflow may meet the needs of many users but not all. Some users will have unique needs that require additional evaluation for either new metrics or other variables that may only apply in particular circumstances. Further, not all metrics and diagnostics would be appropriate for all data products. For example, systematic process-based evaluation of statistically downscaled products (such as metrics based on feature tracking) is often not feasible due to the limited variety or temporal frequency of variables.

Discussions also touched upon the potential for user-oriented evaluation tools, which reconcile the need for standardization with the diverse applications of downscaled products, and the associated need for maximizing utility and ensuring distinctiveness in metrics collections (as discussed in, e.g., Reed et al. (2022)). One framework for such tools is the Coordinated Model Evaluation Capabilities (CMEC) effort currently underway at DOE, which provides a decentralized framework for sharing of evaluation capabilities.

With the development of such a standard framework, questions arose regarding the establishment of an independent entity responsible for such evaluations and what is required to make that entity viable. Such an independent entity would only function with dedicated effort from contributors, and its continuity could be challenged by funding constraints. In particular, the entity would need dedicated, long-term efforts by contributors to build, support, maintain, and adapt software, as well as dedicated efforts by contributors to ensure robust, rigorous, state-of-the-science, and state-of-the-practice DRCDP evaluation. Given the success of the CMIP and CORDEX examples over decades, many of the pathways necessary to develop such an entity have been established.

4.4 Inherited Biases from GCMs

The biases that GCMs exhibit in their historical simulations relative to historical observations have been very persistent across phases of CMIP and are a thorny issue (Ehret et al. 2012). These biases pose a particular problem for dynamical downscaling techniques because the process modeling that underlies such downscaling is sensitive to biases in initial and boundary conditions (Rummukainen 2016). Also, these biases cannot be corrected without incurring uncertainty, and potentially introducing physically unjustifiable inconsistencies among variables. GCM biases are not likely to be resolved at the regional level in the short-term (i.e., in CMIP6Plus or CMIP7), because the contributions to those biases are complex and emerge from GCM parameterization errors, GCM structural errors, and the internal variability of the Earth system. Anyway, GCM development should seek to improve overall GCM performance and consistency with the physical system, rather than only focusing on removing biases in one particular region. From an end-user standpoint, additional research is needed on whether or not these biases materially impact projections (and subsequent decisions) or if they are relatively benign. From the standpoint of data producers, quantification of biases is essential, and further research is needed on the best ways to address these biases without fundamentally altering the processes or physically based relationships that exist between variables. Evaluation-based approaches for down-selection of GCMs may be desirable when selecting GCMs for particular regions (Goldenson et al. 2023). Further, thoughtful approaches (e.g., Risser et al. 2024, Rahimi et al. 2024) that provide clear justification and uncertainty quantification for the use or the avoidance of bias correction are needed.

4.5 Data Averaging and Weighting

With multiple DRCDPs derived from different GCMs, different ensemble members of a GCM, and different downscaling techniques, the distillation of information for the end-user is especially important and requires caution. Ensembling, which refers to combining simulations with similar physical representations and software architectures into ensembles, has long been a practice of the CMIP program to (1) avoid over-reliance on a single projection, (2) capture a range of climate projections using different GCMs, and (3) recognize that natural variability in the Earth system requires probabilistic projections. While there were many fewer projections available to NCA5 than were available for CMIP6, NCA5 adopted an analogous approach to ensembling as CMIP6 and blended LOCA2 and STAR-ESDM. With the advent of numerous downscaled CMIP6 products, combining them to capture a more comprehensive spread of uncertainty may also yield, depending on how it is performed, a result that has reduced extreme-value information relative to individual DRCDPs, if only the mean is provided. It is important to recognize that the spread in projections across climate products and their ensemble members may contain important information about the uncertainty of future change; however, to date little research has been done on the best way to combine information from multiple methods of data production. This is especially the case for consideration of combining statistical and dynamical products, as well as combining different dynamical products whose domains overlap with the geography of interest. Advances in the application of meta-analysis to boost statistical significance, such as recent work to constrain equilibrium climate sensitivity (Sherwood et al. 2020), would be useful in boosting confidence in projections.

The consideration of similarities in base models, whether referring to GCMs or downscaled products, is a complex aspect of evaluating downscaled products. Determining the weights of ensemble projections, particularly when dealing with outliers, requires a careful analysis of model lineage and output similarity (Pennell and Reichler 2011). Determining whether an outlier is signaling a real climatic possibility or is simply the result of issues faced by a particular model remains a key challenge for climate scientists and requires exhaustive evaluation of the outlier.

4.6 Metrics and Diagnostics

Metrics and diagnostics are measures and depictions of differences between a model product and a reference data set, and are fundamental for climate data evaluation. Workshop participants highlighted many examples where new metrics and diagnostics research could address questions related to whether climate data products (individually or collectively) are fit for purpose when employed in particular use cases.

Beyond temperature and precipitation: Discussions among participants revealed a consensus that evaluation metrics should extend beyond traditional measures of surface air temperature and precipitation. This includes development of metrics and reference products for wind, snow, soil moisture, runoff, circulation, humidity, evapotranspiration, and radiation. Additionally, metrics are needed to relate atmospheric (e.g., mid-to-upper troposphere) and surface states (e.g., sea surface temperatures) to surface air temperature, precipitation, and sub-daily extremes. Such metrics are integral to a more holistic understanding of key processes governing regional climate. However, many statistically downscaled data sets currently do not provide data to support such additional evaluation, and the lack of sufficient training data complicates such extended evaluation. This suggests a need to develop hybrid approaches (as discussed in chapter 2) or new observational products.

Feature-based metrics: Feature-based metrics are crucial for understanding the spatial and temporal characteristics of weather and climate features in projections, particularly for those features that have significant implications for regional impacts. Such metrics should encompass surface air temperature features (e.g., heat domes), contiguous regions of precipitation, atmospheric rivers, low- and high-pressure systems, and mesoscale convective systems as essential elements of a comprehensive evaluation strategy.

Intervariable relationships: A significant gap exists in the evaluation of the relationships between different atmospheric and surface variables (e.g., the tendency for wind speeds to be lower when temperatures are high). Evaluating DRCDPs based on these relationships is imperative for accurately modeling and predicting climate behavior for two reasons: (1) because of the strong connection between variable relationships and impact-relevant climatic processes and (2) for data users to have confidence that DRCDP data sets can provide useful information on those intervariable relationships that represent compound risks.

Modes of variability: Better metrics are needed that capture teleconnections between large-scale interannual-to-multi-decadal modes of variability in the climate system and regional climate. One example is to evaluate the relationship between sea surface temperatures (SSTs) and precipitation patterns. This relationship is critical for accurately modeling climate dynamics and physics, especially in regions where El Niño and La Niña are major drivers of surface temperature and precipitation.

Complex indices from external communities: The need to include more complex indices defined by user communities such as fire weather indices and drought conditions was emphasized in discussions. Such metrics are increasingly important in the context of climate change, where the frequency and intensity of multi-factor extreme events are changing.

Weather types: The ability to distinguish and accurately evaluate the skill of a DRCDP for different weather types is an important step to more broadly characterize the skill of a dynamical downscaling product. For example, convective and stratiform rainfall have distinct impacts on regional hydrology and ecosystems and are therefore crucial for climate impact assessments, but metrics for disentangling different precipitating storm types remain a challenge. Extending the evaluation of DRCDPs to encompass a broader range of weather events enables a more robust characterization of a downscaling method's effectiveness and its implications for regional climate predictions.

Focus on user needs: The workshop emphasized the importance of developing metrics inspired by user needs and applications. These use-inspired metrics would allow for the assessment of climate data in the context of real-world applications and sector-specific requirements.

Broader spatial coverage: The geographical scope of evaluation needs to expand beyond the CONUS to at least cover the OCONUS (Basile et al. 2024). Expanding the evaluation of climate projections for a broader range of regions is crucial for understanding climate impacts on a global scale and for supporting resilience and adaptation in regions that may be underrepresented in current regional climate projections.

The role of resolution: There is a need for a more nuanced understanding of how spatial and temporal resolution affects the quality of climate data products. The value introduced by adding spatial resolution to a climate projection (e.g., going from 100km to 25km or 25km to 4km) must be demonstrated, balancing both the greatly increased computational demands and time required to produce high-resolution data against the enhanced detail and potential improvements in capturing relevant climatic processes. This is especially true at even finer spatial resolution (~1km) where the ability to ground-truth the product against observations is extremely limited. For statistically downscaled products, which rely completely on the veracity of gridded historical data, comparisons of different options reveals the difficulty of interpolating between weather stations (Behnke et al. 2016, Ullrich 2023, Walton and Hall 2018) needed to achieve high spatial resolution. Selecting any downscaled data set based largely on high spatial resolution, without considering the realism of the actual variables used for impact analysis, should be carefully considered.

Evaluation of regional climate sensitivity: Finally, the goal of climate downscaling methods is to aid in the evaluation of future climate impacts on adaptation decisions. Metrics need to be developed to evaluate the reliability of future projections from different downscaling methods. Evaluation with respect to historical trends, perfect model experiments, or historical climate modes are steps down this path, and a common framework for such evaluation would greatly benefit the entire community. For instance, in the case of the 5th National Climate Assessment (NCA 5 report), the metric used to evaluate and weight models was Equilibrium Climate Sensitivity (ECS), which captures a model's temperature sensitivity to increase atmospheric CO_2 levels (Massoud et al. 2023). While this selection seems reasonable, as pointed out in Pierce et al. (2009) there may be little relationship between ECS and regional model performance. Additional metrics are likely needed for regional scales, such as hydrologic sensitivity or magnitude (and sign) of precipitation change, often normalized by warming.

Notably, there was broad consensus for the establishment of a process for selecting valuable metrics, as suggested in Reed et al. (2022). The chosen metrics should be capable of guiding the evaluation of statistically and dynamically downscaled products, as well as hybrid approaches.

4.7 Other Gaps and Challenges in DRCDP Evaluation

The workshop emphasized a significant challenge in the field of climate science: effectively linking the evaluation of climate models and DRCDPs with the characterization and communication of their confidence and credibility. There is a noted danger in assessing or establishing model credibility based entirely, or partially, on its use in applications or when its usage is mandated by policy directives. This underscores the need for independent evaluation to provide unbiased, rational, scientifically based assessments of downscaled products.

However, there are barriers to conducting these assessments and the workshop highlighted several opportunities for advancing the science and applicability of data evaluation, which we summarize here.

Evaluating techniques versus output: When evaluating downscaling methods, it is vital to distinguish between evaluating downscaling techniques themselves and the output they produce, because the former includes the implementation of methods and operational choices, while the latter also evaluates the selection of training and/or evaluation data. For example, one common method for evaluating statistical techniques involves downscaling coarsened observations and comparing the results to the original high-resolution data set. While this method provides some insights into the sub-grid information that the statistical method introduces, along with nonlinearities that need to be resolved to avoid grid-scale biases, it fails to address the significant set of biases introduced by global climate models. Therefore, this method should be considered as part of a broader, more inclusive evaluation strategy rather than as a standalone solution.

Evaluating performance in an uncertain future: Understanding the relationship between historical model performance and future projections remains a fundamental challenge. There is a need for a systematic approach to analyze how biases in historical simulations are related to future projections, including the credibility of those projections. This task is made more complex by the differences in regulatory frameworks among various data users, especially where the use of bias-corrected data can be both a limitation and a necessity.

Quantification of uncertainty: Another concern raised in the workshop is the quantification of uncertainty in both observational and downscaled products. An in-depth investigation into the uncertainties specific to each downscaling approach is also needed. Communicating this understanding of uncertainties to end-users in a manner that informs their selection and use of downscaled products remains a considerable challenge.

Guidance on translating these uncertainties into impact models is essential for robust decision-making.

New high-quality observational products: The availability and quality of observational data are critical to model evaluations. The workshop highlighted the need for more observations in data-sparse regions and for variables that are currently not widely available. Furthermore, ensuring that observations are publicly available is important for comprehensive evaluations. Observational data products, particularly their methodology for processing observations, require rigorous evaluation themselves to ensure their suitability for evaluating DRCDPs, a concern that becomes even more prominent outside of CONUS, where observations are relatively sparse and options for data products may be severely limited.

Effect of spread in observational products: Gridded observational products can vary substantially depending on the measurements used and the methods for processing those data. However, evaluation efforts rarely consider multiple observational products or account for this spread in observational data products. More research is imperative to estimate spread in observational data products, and the potential impact the choice of observational product has on the conclusions of that evaluation.

Selecting a subset of products: Specific criteria for the selection of a particular product or products for use in a study or report are inevitably tied to the specific decision context. However, identifying general "best practices" for the selection of products based on scientifically justifiable choices was a common refrain during the workshop to help in guiding groups who lack capacity to keep abreast of this fast-moving field. This is a common practice for all mature professions and timely for the climate adaptation field. However, defining such criteria would require substantive discussions between producers, evaluators, and users, with the added benefit that such discussions would be helpful in developing a better understanding of the whole ecosystem of products.

4.8 References

Ahn, MS, PJ Gleckler, J Lee, AG Pendergrass, and C Jakob. 2022. "Benchmarking simulated precipitation variability amplitude across time scales." *Journal of Climate* 35(20): 6773–6796, https://doi.org/10.1175/JCLI-D-21-0542.1

Basile, S, CW Avery, A Grade, and AR Crimmins. 2024. "To be policy-relevant, future climate research must include the noncontiguous United States." *Proceedings of the National Academy of Sciences of the United States of America* 121(13): e2315505121, https://doi.org/10.1073/pnas.2315505121

Behnke, R, S Vavrus, A Allstadt, T Albright, WE Thogmartin, and VC Radeloff. 2016. "Evaluation of downscaled, gridded climate data for the conterminous United States." *Ecological Applications* 26(5): 1338–1351, https://doi.org/10.1002/15-1061

Bock, L, A Lauer, M Schlund, M Barreiro, N Bellouin, C Jones, GA Meehl, V Predoi, MJ, Roberts, and V Eyring. 2020. "Quantifying progress across different CMIP phases with the ESMValTool." *Journal of Geophysical Research – Atmospheres* 125(21): e2019JD032321, https://doi.org/10.1029/2019JD032321

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

Lee, J, PJ Gleckler, MS Ahn, A Ordonez, PA Ullrich, KR Sperber, KE Taylor, YY Planton, E Guilyardi, P Durack, and C Bonfils. 2023. "Objective Evaluation of Earth System Models: PCMDI Metrics Package (PMP) version 3." *Geoscientific Model Development* 17: 3919–3948, https://doi.org/10.5194/gmd-17-3919-2024

Massoud, EC, HK Lee, A Terando, and M Wehner. 2023. "Bayesian weighting of climate models based on climate sensitivity." *Communications Earth & Environment* 4(1): 365, https://doi.org/10.1038/s43247-023-01009-8

Pennell, C, and T Reichler. 2011. "On the effective number of climate models." *Journal of Climate* 24(9): 2358–2367, https://doi.org/10.1175/2010JCLI3814.1

Pierce, DW, TP Barnett, BD Santer, and PJ Gleckler. 2009. "Selecting global climate models for regional climate change studies." *Proceedings of the National Academy of Sciences of the United States of America* 106(21): 8441–8446, https://doi.org/10.1073/pnas.0900094106

Potter, GL, DC Bader, M Riches, A Bamzai, and R Joseph. 2011. "Celebrating two decades of the Program for Climate Model Diagnosis and Intercomparison." *Bulletin of the American Meteorological Society* 92(5): 629–631, https://doi.org/10.1175/2011BAMS3018.1

Rahimi, S, L Huang, J Norris, A Hall, N Goldenson, M Risser, DR Feldman, ZJ Lebo, E Dennis, and C Thackeray. 2024. "Understanding the Cascade: Removing GCM biases improves dynamically downscaled climate projections." *Geophysical Research Letters* 51(9): e2023GL106264, https://doi.org/10.1029/2023GL106264

Reed, KA, N Goldenson, R Grotjahn, WJ Gutowski, K Jagannathan, AD Jones, LR Leung, SA McGinnis, SC Pryor, AK Srivastava, PA Ullrich, and CM Zarzycki. 2022. "Metrics as tools for bridging climate science and applications." *Wiley Interdisciplinary Reviews: Climate Change* 13(6): e799, https://doi.org/10.1002/wcc.799

Righi, M, B Andela, V Eyring, A Lauer, V Predoi, M Schlund, J Vegas-Regidor, L Bock, B Brötz, L de Mora, F Diblen, L Dreyer, N Drost, P Earnshaw, B Hassler, N Koldunov, B Little, SL Tomas, and K Zimmerman. 2020. "Earth System model evaluation tool (ESMValTool) v2.0– technical overview." *Geoscientific Model Development* 13(3): 1179–1199, https://doi.org/10.5194/gmd-13-1179-2020

Risser, MD, S Rahimi, N Goldenson, A Hall, ZJ Lebo, and DR Feldman. 2024. "Is bias correction in dynamical downscaling defensible?" *Geophysical Research Letters* 51(10): e2023GL105979, https://doi.org/10.1029/2023GL105979

Rummukainen, M. 2016. "Added value in regional climate modeling." *Wiley Interdisciplinary Reviews: Climate Change* 7(1): 145–159, https://doi.org/10.1002/wcc.378

Sherwood, SC, MJ Webb, JD Annan, KC Armour, PM Forster, JC Hargreaves, G Hegerl, SA Klein, KD Marvel, EJ Rohling, M Watanabe, T Andrews, P Braconnot, CS Bretherton, GL Foster, Z Hausfather, AS von der Heydt, R Knutti, T Mauritsen, JR Norris, C Proistosescu, M Rugenstein, GA Schmidt, KB Tokarska, and MD Zelinka. 2020. "An assessment of Earth's climate sensitivity using multiple lines of evidence." *Reviews of Geophysics* 58(4): e2019RG000678, https://doi.org/10.1029/2019RG000678

Ullrich, PA. 2023. Validation of LOCA2 and STAR-ESDM Statistically Downscaled Products. LLNL-TR-856094: 1085545, https://doi.org/10.2172/2202926

Walton, D, and A Hall. 2018. "An assessment of high-resolution gridded temperature datasets over California." *Journal of Climate* 31(10): 3789–3810, https://doi.org/10.1175/JCLI-D-17-0410.1

5.0 Technical Talks

Lead Authors:

Daniel Feldman, Lawrence Berkeley National Laboratory, drfeldman@lbl.gov

Paul Ullrich, Lawrence Livermore National Laboratory, ullrich4@llnl.gov

Contributing Authors:

Annarita Mariotti, National Oceanic and Atmospheric Administration HQ annarita.mariotti@noaa.gov

Elias Massoud, Oak Ridge National Laboratory massoudec@ornl.gov

Hugo Lee, NASA Jet Propulsion Laboratory huikyo.lee@jpl.nasa.gov

Alex Hall, University of California, Los Angeles alexhall@ucla.edu

Muge Komurcu, NASA Ames muge.komurcu@nasa.gov

Ruby Leung, USDOE Pacific Northwest National Laboratory ruby.leung@pnnl.gov

Melissa Bukovsky, University of Wyoming melissa.bukovsky@uwyo.edu

Greg Tierney, North Carolina State Climate Office getierne@ncsu.edu

Owen Doherty, Eagle Rock Analytics owen@eaglerockanalytics.com

5.1 Chapter Summary

The workshop organizers solicited brief technical talks from the attendees to discuss the current state of the practice and state of the science. These talks also provided a diversity of perspectives on the biggest issues facing decision-relevant climate data products, and potential future directions to tackle these issues. This chapter summarizes the nine presentations given at the workshop.

5.2 Envisioning the Next Generation of U.S. Climate Predictions and Projections

Annarita Mariotti, NOAA HQ

A national strategy for next-generation U.S. climate predictions and projections is crucial to meet the needs for decision-relevant climate products. The development of this strategy starts with the realization that significant accelerated progress is both necessary and possible over the next several years, based on the scientific and technological opportunities now at hand. Fundamental to the strategy are the strengths of the U.S. climate modeling community, including innovation and diversity in the federally funded climate modeling groups, academia, and the private sector. A strategy is to encompass unprecedented levels of coordination and resources for transformative opportunities in both science and science-to-service pathways, and the enabling environments necessary for progress. These include advances in computing infrastructure for climate modeling, data storage and data analytics to support both science and service needs, and workforce and partnerships. Given the increasing diversity of modeling types supporting decisions, the strategy would include coordination for the development of a model-agnostic evaluation framework based on both scientific understanding and decision-relevant metrics.

5.3 Model Weighting Based on Equilibrium Climate Sensitivity

Elias Massoud, Oak Ridge National Laboratory

During a recent workshop, the use of Bayesian Model Averaging (BMA) on CMIP6 models was discussed, with the aim of constraining the model ensemble based on ECS. The BMA method allows for combining information from multiple models, providing a more detailed understanding of climate projections. In this context, BMA helped address the 'hot model' problem that is known in CMIP6, where some models show higher ECS values than others, potentially skewing overall climate projections. The weights derived from BMA in this work were used in the recent NCA 5th report, where they were applied to downscaled models used in regional climate assessments. This approach offers a realistic and probabilistic assessment of future climate scenarios on a local scale, aiding regional decision-making.

5.4 A Multi-Resolution Framework for Evaluating High-Resolution Climate Simulations

Kyo Lee, NASA Jet Propulsion Laboratory (JPL)

This presentation explored evaluation of effective data resolution by developing a methodology to compare spatial patterns from various sources using Hierarchical Equal Area isoLatitude Pixelization (HEALPix) within JPL's Regional Climate Model Evaluation System. This work focused on temperature trends and mean precipitation, using HEALPix's equal-area pixels and hierarchical structure to enable efficient remapping and comparison of data sets at different resolutions. A hierarchical data analyzer (HDA) has been designed to support multi-resolution analysis and the examination of spatial variance, map differences, and anisotropic patterns. This analysis technique was demonstrated by examining temperature trends in the Northwest region and comparing precipitation spatial variances between DOE's regionally refined model and satellite observations. The findings highlight the added value of higher spatial resolution in understanding spatial variability, though the 25-km resolution may not capture small-scale precipitation processes accurately. This framework can assess the value of high spatial resolutions in other downscaled data sets.

5.5 A Seamless Approach for Evaluating Climate Models across Spatial Scales/Evaluating Impact of Bias Correction on Downscaling Uncertainty

Alex Hall, University of California Los Angeles

Issues surrounding bias correction pose significant challenges for decision-relevant climate data products. Bias correction is often necessary for decision relevance, yet its effect on climate change signals is largely unevaluated. This creates a significant (and unaccounted for) source of uncertainty. To begin to address this, dynamically downscaled simulations of future climate over the western U.S. were compared with and without a prior bias correction of the driving GCM data. It was found that the impacts of bias correction on regional temperature and precipitation change signals are small relative to other uncertainty sources such as internal variability and GCM diversity. Furthermore, it was found that in some cases (e.g., snow projections) bias correction produces a more physically credible solution. Much work remains to understand the impacts of bias correction, including why bias correction distorts regional signals, and the full spectrum of circumstances when bias correction actually produces more physically defensible results.

5.6 Assessing Physical Climate Risks: Challenges and Opportunities

Muge Komurcu, NASA Ames

Dynamical downscaling using convection-permitting regional climate modeling (<=4km horizontal resolution) can improve both mean and extreme features of climate in local scales compared to the driver Earth system model (or climate reanalysis) as shown in Komurcu et al. 2018. However, it is computationally expensive to run these models, and to post-process, store, and distribute the resulting data, which makes it challenging to downscale multiple models, scenarios, and realizations, to assess uncertainties and probabilities of projected change. Though missed by many published studies, it is also essential to downscale a climate reanalysis data set for the region using the same model set-up first to evaluate the methodology prior to applying it to downscale ESM historical climate and projections. Resulting higher-resolution projections enable many opportunities for collaborations between climate scientists and other experts (e.g., architects, civil engineers, hydrologists, economists) to assess physical climate risks, to produce region-specific adaptation pathways, and to create climate-resilient urban and infrastructure design. However, dedicated funding opportunities for these collaborations are limited. Public and stakeholder outreach undertaken by physical scientists can be crucial to promote these opportunities.

5.7 Designing and Using Heterogeneous Ensembles of Climate Scenarios for Decision-Making

L. Ruby Leung and Claudia Tebaldi, Pacific Northwest National Laboratory

The impacts of extreme weather risks vary with the likelihood of extreme weather events and their consequences. Consequences are largely driven by the intensity of weather events for which model resolution matters. To estimate the likelihood of extreme weather events in specific regions, large ensembles of simulations are critical because of the significant uncertainty in large-scale circulation. Hence to support decision-making regarding the impacts of extreme weather risks, large ensemble size, high resolution, and multiple models are essential elements of ensemble modeling for robust estimation. Despite the groundbreaking performance of the Simple Cloud-Resolving Energy Exascale Earth System Model's Atmosphere Model (SCREAM) global cloud-resolving model on the Frontier exascale computer (1 SYPD), the type of ensembles needed to quantify extreme weather risks and their impacts is unattainable without transformational expansion of computing resources and improvements in computational performance. Currently, a heterogeneous ensemble of simulations consisting of regional and global models at different resolutions with different ensemble sizes exist. Challenges for the community include: (1) how to combine strengths from existing small-ensemble/single-member higher-resolution simulations and medium/large ensemble of lower-resolution simulations; (2) how to design a new heterogeneous ensemble through coordinated efforts; and (3) how to augment physical modeling with machine learning (e.g., ensemble boosting) and leverage the heterogeneous ensemble to estimate the probability of extreme events.

5.8 The North American Coordinated Regional Climate Downscaling Experiment (NA-CORDEX): Overview and Status

Melissa Bukovsky, University of Wyoming

Rachel McCrary, National Center for Atmospheric Research

Dominique Paquin, Ouranos

Christopher McCray, Ouranos

William Gutowski, Iowa State University

Sara Pryor, Cornell University

The CORDEX vision is to advance and coordinate the science and application of regional climate downscaling through global partnerships. It emerged from a 2009 World Climate Research Programme (WCRP) call for enhanced downscaling coordination and serves as a formal CMIP diagnostic MIP. Coordination enhances the functionality of model ensembles for end-users and makes optimum use of a limited resource pool. However, achieving coordination has proven challenging due to the diverse array of disparate downscaling endeavors pursued under varying rationales, compounded by inadequate levels of support. Data storage also continues to be a challenge. NA-CORDEX, despite these challenges, has to date served a wide audience. CMIP5 downscaling efforts have garnered 160+ data set and 100+ journal citations, contributed to three IPCC reports and the Interactive Atlas, with the 40TB public archive maintaining a monthly download rate of about 5.1TB since completion. The coordination and production of downscaled CMIP6 simulations is underway, but it is expected to face similar challenges. For international CORDEX and protocol details, visit cordex.org, and for North American efforts, visit na-cordex.org.

5.9 A Unified Framework for Evaluating the Risk of Extreme Temperature Events

Greg Tierney, North Carolina State Climate Office

As heat is the leading cause of weather-related deaths in the United States, risk assessment of extreme temperature events is crucial in understanding the potential impacts of a changing climate on human health. By employing an intensity-duration-frequency (IDF) curve framework, long used in the hydrological community, a concise but comprehensive analysis can be conducted for several return periods and event durations simultaneously. Applying this framework to future climate data products is beneficial along two paths. First, the IDF analysis can be a diagnostic tool evaluating the model's ability to realistically replicate extreme temperature events at an appropriate magnitude and frequency. Second, IDF analysis can be used prognostically to evaluate the future risk of extreme temperature events, enabling communities to better prepare future adaptation and mitigation measures. Expansion of this framework beyond temperature to broader human health indices (when possible, given the underlying climate data products) will serve to enrich such risk assessment going forward.

5.10 Making Complex Data Actionable for Regional Decision-Making via the Cal-Adapt: Analytics Engine

Owen Doherty, Eagle Rock Analytics

Making unapproachably large quantities of climate projections accessible and useful for decision-makers requires an understanding of how the data will be used, scientifically informed analysis, appropriate statistical techniques, and computational/technological advances that empower end-users to use the climate projections. In turn, developing such a solution requires collaborative contributions, innovation, and a willingness to iterate from engaged and empowered users, social scientists, climate scientists, and computer/data scientists. Through co-production at Cal-Adapt we have assembled a collective of energy sector experts and leaders, government regulators, social and climate scientists across national laboratories, the University of California, and industry, as well as technical innovators including staff from Amazon Web Services to collectively produce the Cal-Adapt: Analytics Engine. The Cal-Adapt: Analytics Engine provides the data, compute resources, and expert guided notebooks to enable users to execute complex analysis that lead to regional decision-making using scientifically and statistically rigorous workflows. In this presentation we walk through a number of examples, including use of observational data at weather stations for on-the-fly bias correction to create future time series for energy sector stakeholders to use in energy system models and vulnerability assessments. We also demonstrate how global warming levels can be used in vulnerability assessments, and highlight the need for guidance and best practices to support users to enable decision-making at the regional scale.

6.0 Summary and Future Vision

Lead Authors:

Daniel Feldman, Lawrence Berkeley National Laboratory, drfeldman@lbl.gov

Paul Ullrich, Lawrence Livermore National Laboratory, ullrich4@llnl.gov

Contributing Authors:

Melissa Bukovsky, University of Wyoming, melissa.bukovsky@uwyo.edu

Ethan Gutmann, National Center for Atmospheric Research, gutmann@ucar.edu

Kripa Jagannathan, Lawrence Berkeley National Laboratory, KAJagannathan@lbl.gov

Andy Jones, Lawrence Berkeley National Laboratory, adjones@lbl.gov

Ruby Leung, Pacific Northwest National Laboratory, ruby.leung@pnnl.gov

Stefan Rahimi, University of Wyoming, srahimi@uwyo.edu

Tanya Spero, U.S. Environmental Protection Agency, spero.tanya@epa.gov

Adam Terando, U.S. Geological Survey, aterando@usgs.gov

6.1 Chapter Summary

The Decision-Relevant Climate Data Product workshop sought to identify ongoing gaps in connecting the state of the science of localized climate projections to actionable climate information for end-users. This chapter summarizes these gaps, describes lessons learned from the global climate modeling community with the Coupled Model Intercomparison Project, and presents near- and medium-term upcoming challenges and opportunities. The chapter concludes by describing the features of a viable community of practice for DRCDPs so that the large and growing body of climate data users has salient, credible, authoritative, and accessible regional climate projection information.

6.2 Background

This workshop arose due to a set of ongoing, as well as a few acute, needs for decision-relevant climate information across the broad ecosystem of data producers, translators, evaluators, and end-users. Presentations and discussions focused on sharing of efforts currently underway and identifying gaps that limit the utility of climate data products by researchers and end-users.

6.3 Summary of Identified Gaps

Identified gaps were organized in three categories, reflecting the three themes of the workshop: data production, data use, and data evaluation. In addition, workshop discussions identified two cross-cutting gaps related to observational data and cyberinfrastructure.

6.3.1 Gaps in Data Production

Workshop discussions highlighted data production needs that include: (1) more high-temporal-resolution (i.e., sub-daily to sub-hourly) climate projections to address climate change impacts on renewable energy and extreme weather; (2) quantified uncertainties from the methods used for generating climate projections; (3) better methods for diagnosing variables relevant for climate change impacts such as flood extent, soil moisture, and wildfires; and (4) data homogenization and standardization to improve data availability and utility across workflows. From a data access and availability perspective, we noted the need for DRCDPs to follow findability, accessibility, interoperability, and reusability (FAIR) principles. This is because of a lack of community standards for variable names and file metadata that preclude interoperability of these data sets, and the lack of a catalog of decision-relevant data products that would allow end-users to learn about what is available. An improved organizational structure will reduce redundancy in data production Additionally, the continuing, pressing need for high-resolution climate projections outside the CONUS, particularly over islands and in the Arctic, was expressed; however, this need is underscored by the lack of high-quality observational data in these regions, particularly data available with broad spatial coverage and long duration.

6.3.2 Gaps in Data Use

Data users identified an outstanding need for decision-relevant variables and increased spatial and temporal resolution, along with tools and computing resources that allow decision-makers to easily work with large amounts of climate data. Data users often mention the lack of specific and operational guidance on how to (and how not to) use climate data for specific applications. A dire need for scientifically rigorous and decision-relevant credibility evaluations of data products was also identified as a key gap, particularly as there is a proliferation of many different types of climate data products. As an overarching issue, there is often limited engagement and inclusion of user perspectives in climate data production and evaluation. This has led to many gaps in understanding how climate data are eventually used by different types of decision-makers and users in different types of applications. With the NCA6 now starting to spin up, there is a need for formal guidance (developed through collaborative engagements between data producers, users, and evaluators) on how the federal government should identify and select climate data products so as to avoid ad hoc solutions.

6.3.3 Gaps in Data Evaluation

The workshop identified a clear and outstanding need for a standard framework for climate data product evaluation, with the caveat that some flexibility is required to incorporate differences in user needs. Discussions further identified a need for **new metrics beyond temperature and precipitation**, particularly key impact variables such as hub-height winds, snow, humidity, and radiation. Future work could also consider relationships and covariances between variables, as well as more complex secondary variables such as fire weather conditions or drought indices. New metrics should also **incorporate observational uncertainties** when scoring data products, particularly in observation-poor regions. Conversations also highlighted potential benefits from an independent organization for evaluating climate data products, particularly to build trust in these data products and support consistent evaluation across products.

6.3.4 Cross-Cutting Gaps: Observational Data

Observational data products are the basis for testing regional climate projections, and yet there are many choices made in constructing these products (e.g., bilinear interpolation between stations to form a gridded product, homogenization and sanitization of underlying observations, movement of stations, data-filling procedures, etc.) that materially impact the comparisons with downscaling projections. Major gaps exist in terms of research that establishes the fitness for purpose (or lack thereof) of observational products for generally establishing confidence in projections and more specifically establishing confidence in such projections for specific climate impacts (e.g., gridded observational products are inappropriate to test a projection's extreme precipitation performance unless the gridding procedure preserves precipitation extremes such as Risser et al. 2019. Research is needed to establish a chain of custody between observations, observational products, and their usage for DRCDP evaluation, along with improved quantification of observational uncertainty. And, as mentioned earlier, there is a need for more observational products outside CONUS.

6.3.5 Cross-Cutting Gaps: Cyberinfrastructure

A recent President's Council of Advisors on Science and Technology (PCAST) report on extreme weather risk in a changing climate similarly emphasized the need for quantification of risk from extreme weather, the maintenance of an extreme weather data portal, and the need to inventory and release relevant federal data (PCAST 2023). Indeed, presently available cyberinfrastructure is largely insufficient for distribution of decision-relevant climate data sets. Because decision-relevant climate data products have significantly higher resolution and (in some cases) a greater number of available variables, these data sets, when taken altogether, have grown to be similar in magnitude to the size of the CMIP6 archive. The workshop highlighted the need to investigate partnerships with federated climate data infrastructure like the Earth System Grid Federation (ESGF) or cloud service providers. Additionally, there is a need for **software and hardware to support no-cost-to-user server-side analysis and evaluation of data products so as to sidestep the need for bulk data transfer and enable greater access and lower barriers to entry for using these data products**.

6.4 Parallels with the Coupled Model Intercomparison Project

Decision-relevant climate information is related to, but distinct from, the international consortium of climate modeling experiments that comprise the CMIP, largely because (1) the information contained in CMIP models is too coarse to be used for local decisions and (2) global climate modeling is fundamentally an academic undertaking to advance the scientific understanding of the Earth system. CMIP models and experiments are sometimes aligned with, and sometimes separate from, the requirements for salient, credible, authoritative, and accessible regional climate projection information.

However, there are obvious parallels between the gaps identified above and the efforts that have been undertaken as part of CMIP. For instance, data distribution for CMIP is largely handled through the ESGF (Cinquini et al. 2014), a resource that allows essentially any user to query and download available data sets. For data to be contributed to ESGF, groups are required to make it compliant with Climate and Forecast (CF) standards (Hassell 2017, Eaton et al. 2024), which prescribes specific variable naming conventions and file metadata. To be considered part of CMIP, modeling groups must also submit a set of common model simulations known as the Diagnostic, Evaluation and Characterization of Klima (DECK) (Eyring et al. 2016). These simulations, in conjunction with common standards for data format, permit any group to perform intercomparisons using their own evaluation tools. Consequently, groups such as PCMDI are able to provide interactive graphics to intercompare CMIP model performance on a variety of relevant metrics. Given that the standards and approach to climate model intercomparison are less than 30 years old (Touzé-Peiffer et al. 2020), lessons learned from CMIP thus provide a potential foundation for analogous efforts within the decision-relevant climate data community.

6.5 Upcoming Challenges and Opportunities

Conversations at the workshop further highlighted upcoming challenges and opportunities for the decision-relevant climate data, particularly collaborations with active groups that produce or rely on climate data.

The Sixth National Climate Assessment (NCA6): The NCA is a congressionally mandated report on the state of the climate system, with the sixth report expected in 2027. The NCA5, published in November 2023, used two statistically downscaled products for historical climate and future projections. These statistical ensembles reflected advances in techniques that had been used in the fourth report, and they were the only available datasets during the development of NCA5. These products were eventually averaged to produce a single projection of future change, despite such a step being scientifically questionable (Zwiers et al. 2013). Efforts to validate these data sets kicked off in early 2023 (Ullrich 2023), despite acknowledgement that a process should have been in place earlier. For the next NCA6, several new and/or updated statistically and dynamically downscaled data products are expected to meet the basic data and publication needs of this report. There will be a need to develop formal and objective criteria for selecting or weighing available products to maximize utility for data users.

CMIP6plus and CMIP7: The rapid expansion of global climate model projections, with more modeling centers contributing to CMIP experiments, more ensemble members per modeling center contribution, and more CMIP experiments, has created (and is contributing to the growth in) barriers to entry for the very stakeholders that the climate projection data is meant to serve. CMIP6 alone has been responsible for the generation of over 15.8 PB of climate data (Balaji et al. 2018), a number that is expected to increase at least fivefold for CMIP7 (Stockhause et al. 2024). Downscaled data sets derived from CMIP ensemble members will further exacerbate data storage requirements, particularly when considering a growing number of impacts-relevant variables. Efforts to steer the generation of decision-relevant climate projection data onto a sustainable pathway that fully harnesses the advancements in climate science that are contained within CMIP, while handling data growth and the distillation of climate information that end-users require, are sorely needed.

North American Coordinated Regional Downscaling EXperiment (NA-CORDEX): CORDEX (cordex.org), including its North American branch (na-cordex.org), has a long history of success in coordinating downscaling efforts, using a community vetted protocol for downscaling (dynamical and statistical) and data archiving to facilitate simulation intercomparison and enhance data usefulness. As such, NA-CORDEX provides an excellent opportunity and platform for coordination, data dissemination, and knowledge exchange with the decision-relevant climate data community. CORDEX CMIP6 downscaling in North America is still in its early stages, leaving ample opportunity to explore and engage in this effort.

Outstanding questions in climate data production and data science: The strengths and weaknesses of statistical and dynamical downscaling methods, and questions about bias correction and training data, will also persist without intervention. Furthermore, the operationalization of non-traditional downscaling methods, especially data-driven approaches, could amplify the problems of enormous data volumes and challenges for climate data access.

Interagency coordination: Among federal agencies, there are both overlapping and unique needs for decision-relevant climate data. However, data production efforts have largely occurred in isolation, an issue that could be addressed through better coordination. Sharing of data products would also allow us to better quantify and constrain uncertainties in future projections.

6.6 Definition of a Community of Practice

The workshop discussions were very frank on the challenges inherent in the establishment of a Community of Practice (COP) for DRCDP. The primary focus of a community of practice would be to facilitate and encourage communication between parties and to leverage limited human and computational resources (see Figure 12). These efforts have been attempted before (e.g., Barsugli et al. 2013), but have not resulted in the establishment of such a community. Past efforts have had difficulty maintaining momentum, in part because of inconsistent agency support and insufficient buy-in from across the community. Recognizing and learning from these efforts, it is clear that more sustained support is required from a diverse community of data producers, evaluators, and users, inevitably tied to the perceived value in the interactions supported through ongoing collaboration.

Potential goals that could be tackled by a COP include:

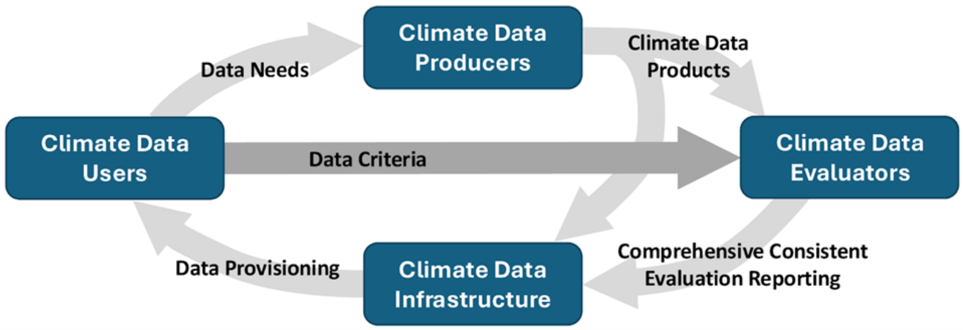


Figure 12. Communication pathways that are needed and potentially addressed by a community of practice.

Regular assessment of the state of the discipline: The COP could regularly assess the state of the discipline, identify emergent gaps, and track progress on filling those gaps. This goal could be accomplished through regular (potentially annual) meetings among members of the COP, structured similarly to this workshop.

Coordinated data production and cyberinfrastructure support: As noted earlier, because of the significant computational cost of downscaling, particularly dynamical downscaling, it is highly desirable to maximize the value from each simulation performed and avoid redundancy in data production. Bringing together data producers and users would support the selection of variables, along with their corresponding spatial and temporal resolution, that maximize utility among user groups. A COP would provide a platform for sharing those needs among relevant groups.

Development of expert guidance: Data users highlighted a pressing need for expert guidance related to the use of DRCDPs. The COP would provide a forum to exchange knowledge on decision-relevant climate data, including how and when particular data products should be used, and where care needs to be taken in their use.

Development of data user engagement and co-production: Engaged data users are critical for COP success because they can ensure that end-user considerations are prioritized in scientific analyses. The COP can, on a rotating basis, undertake specific risk analyses in direct collaboration with the end-users who are participating in the COP and who face those risks. These actions ideally would demonstrate to other end-users the benefits of participating in the COP and encourage additional end-user participation.

Development of standards: The workshop identified a need for common data standards for climate data sets, and the need for a standard evaluation protocol that is cognizant of differences in user need. Committees mandated with the development and maintenance of these standards could be accomplished via a COP.

Sharing experiences and lessons learned: Subgroups within the COP addressing challenges in data production, valuation, and use could meet regularly to discuss shared experiences and lessons learned to enhance the utility of the COP, improve the quality of the data, and better meet the needs of the end users.

Recognition for CoP participation: Ensuring social cohesion, participation, and a sense of community is advanced through regular, meaningful, positive recognition of contributions to the COP and community-building activities. Awards for contributions can, if thoughtfully administered, also incentivize ongoing participation.

Many outstanding questions remain related to the COP, including guidelines for membership, governance structure, mechanisms for agency support, and how to ensure continued relevance and trustworthiness. While the workshop was effective at providing a potential outline for such a COP, further discussions are needed to develop a structure that would support the broader DRCDP community.

6.7 References

Balaji, V, KE Taylor, M Juckes, BN Lawrence, PJ Durack, M Lautenschlager, C Blanton, L Cinquini, S Denvil, M Elkington, F Guglielmo, E Guilyardi, D Hassell, S Kharin, S Kindermann, S Nikonov, A Radhakrishnan, M Stockhause, T Weigel, and D Williams. 2018. "Requirements for a global data infrastructure in support of CMIP6." *Geoscientific Model Development* 11(9): 3659–3680, https://doi.org/10.5194/gmd-11-3659-2018

Barsugli, JJ, G Guentchev, RM Horton, A Wood, LO Mearns, X-Z Liang, JA Winkler, K Dixon, K Hayhoe, RB Rood, L Goddard, A Ray, L Buja, and C Ammann. 2013. "The practitioner's dilemma: How to assess the credibility of downscaled climate projections." *Eos* 94(46): 424–425, https://doi.org/10.1002/2013EO460005

Cinquini, L, D Chrichton, C Mattmann, J Harney, G Shipman, F Wang, R Ananthakrishnan, N Miller, S Denvil, M Morgan, Z Pobre, GM Bell, C Doutriaux, R Drach, D Williams, P Kershaw, S Pascoe, E Gonzalez, S Fiore, and R Schweitzer. 2014. "The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data." *Future Generation Computer Systems* 36: 400–417, https://.doi.org/10.1016/j.future.2013.07.002

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

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

Eaton, B, J Gregory, B Drach, K Taylor, S Hankin, J Caron, R Signell, and P Bentley. 2024. NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.12, https://cfconventions.org/cf-conventions/cf-conventions.pdf

Eyring, V, S Bony, GA Meehl, CA Senior, B Stevens, RJ Stouffer, and KE Taylor. 2016. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization." *Geoscientific Model Development* 9(5): 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016

Hassell, D, J Gregory, J Blower, BN Lawrence, and KE Taylor. 2017. "A data model of the Climate and Forecast metadata conventions (CF-1.6) with a software implementation (cf-python v2.1)." *Geoscientific Model Development* 10(12): 4619–4646, https://doi.org/10.5194/gmd-10-4619-2017

President's Council of Advisors on Science and Technology (PCAST). 2023. Report to the President: Extreme Weather Risk in a Changing Climate: Enhancing prediction and protecting communities, https://www.whitehouse.gov/wp-content/uploads/2023/04/PCAST_Extreme-Weather-Report_April2023.pdf

Risser, MD, CJ Paciorek, TA O'Brien, MF Wehner, and WD Collins. 2019. "A probabilistic gridded product for daily precipitation extremes over the United States." *Climate Dynamics* 53(5): 2517–2538, https://doi.org/10.1007/s00382-019-04636-0

Rohde, RA, and Z Hausfather. 2020. "The Berkeley Earth Land/Ocean Temperature Record." *Earth System Science Data* 12(4): 3469–3479, https://doi.org/10.5194/essd-12-3469-2020

Stockhause, M, M Mizielinski, A Pirani, L Sitz, A Spinuso, M Abdallah, J Baidya, P Durack, and D Ellis. 2024. How do the CMIP7 infrastructure plans support the implementation of the IPCC data guidelines? European Geophysical Union General Assembly 2024. Vienna, Austria. EGU24-17356, https://doi.org/10.5194/egusphere-egu24-17356

Ullrich, PA. 2023. Validation of LOCA2 and STAR-ESDM Statistically Downscaled Products. LLNL-TR-856094: 1085545, https://doi.org/10.2172/2202926

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

Zwiers, FW, LV Alexander, GC Hegerl, TR Knutson, JP Kossin, P Naveau, N Nicholls, C Schär, SI Seneviratne, and X Zhang. 2013. "Climate extremes: challenges in estimating and understanding recent changes in the frequency and intensity of extreme climate and weather events." In Asrar, G, and J Hurrell, Eds., *Climate Science for Serving Society: Research, Modeling and Prediction Priorities*. Dordrecht, Germany. 339–389. Springer, Dordrecht, Netherlands, pp. 339–389, https://doi.org/10.1007/978-94-007-6692-1_13

U.S. Department of ENERGY | Office of Science

Biological and Environmental Research Program