EXECUTIVE SUMMARY

Determining the Temporal and Spatial Scales of Nonstationarity in Temperature and Precipitation across the Continental United States for a Given Emissions Scenario

SERDP Project RC18-1577

DECEMBER 2019

Daniel Feldman
Lawrence Berkeley National Laboratory

Distribution Statement A
This document has been cleared for public release

SERDP
DOD ▪ EPA ▪ DOE

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

EXECUTIVE SUMMARY

Project: RC18-1577

TABLE OF CONTENTS

Page

1.0 INTRODUCTION 1
2.0 OBJECTIVES 3
3.0 TECHNICAL APPROACH 4
4.0 RESULTS AND DISCUSSION 6
5.0 IMPLICATIONS FOR FUTURE RESEARCH AND BENEFITS 9
6.0 LITERATURE CITED 9

LIST OF FIGURES

Page

Figure 1. Flow-chart Indicating How CMIP Model Projections Are Translated into Model Facility-level Risk. 3
Figure 2. Flow Diagram of a Regression-based Analysis. 5
Figure 3. Change in 20-year Daily Return Value of Maximum Surface Air Temperature and Precipitation 7
Figure 4. Differences Between Weather Research and Forecasting (WRF) and LOCA-WRF Projections 8

LIST OF TABLES

Page

Table 1. Changes in Average Daily Precipitation and Maximum Daily Surface Temperature Values (Future-Past) for TAFB 6
Table 2. Changes in Average Daily Precipitation and Maximum Daily Surface Temperature Values (Future-Past) for FBR 6
Table 3. The Ratio of Probability of the Extreme Event in the Future that Corresponded to $p<0.01$ in the Past, and Its Probability for Three Future Periods 7

ACRONYMS AND ABBREVIATIONS

CDF	Cumulative Distribution Function
CMIP	Coupled Model Intercomparison Project
CMIP5	Coupled Model Intercomparison Project – Phase 5
CMIP6	Coupled Model Intercomparison Project – Phase 6
CONUS	Conterminous United States
DEW	Drier Extreme Warming
DoD	Department of Defense
FBR	Fort Bragg
GCM	Global Circulation Model
GEV	Generalized Extreme Value
HighResMIP	High-Resolution Model Intercomparison Project
IPCC	Intergovernmental Panel on Climate Change
LOCA	Localized Constructed Analogs
MLR	Multi-Linear Regression
NCAR	National Center for Atmospheric Research
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RCP4.5	Representative Concentration Pathway, version 4.5
RCP8.5	Representative Concentration Pathway, version 8.5
SERDP	Strategic Environmental Research and Development Program
TAFB	Travis Air Force Base
VIACS	Vulnerability, Impacts, Adaptation, and Climate Services
WRF	Weather Research and Forecasting

ACKNOWLEDGEMENT

The investigating team would like first and foremost to acknowledge the Strategic Environmental Research and Development Program (SERDP) for providing the funding to perform this research, supporting the 2018 Symposium which enabled scientific discussions and presentations, for inviting this team to present and discuss this team's research at the SERDP webinar, and for inviting the team to participate in the Resilient Aviation Infrastructure Workshop. This team would also like to recognize the detailed discussions with hydrologists from the State of California Department of Water Resources including Dr. Michael Anderson, Elissa Lynn, Wyatt Arnold, and Kevin He. Dr. Michael Wehner of the Lawrence Berkeley National Laboratory provided valuable feedback on the Coupled Model Intercomparison Project (CMIP) landscape. The team would also like to recognize the efforts of potential collaborators most especially Dr. Travis O'Brien of the Lawrence Berkeley National Laboratory for attending, discussing, and contributing to the Resilient Aviation Workshop.

Additional discussions with Neil Berg of UCLA and Dave Pierce and Dan Cayan of UCSD were very useful in developing detailed perspectives on the strengths and weaknesses of different downscaling approaches. The team would further like to recognize the guidance of Dr. Jeffrey Arnold of the U.S. Army Corps of Engineers Climate Preparedness and Resiliency Program, since he graciously took time out of his schedule to provide information on how downscaled CMIP data are used to characterize the risk in specific watersheds around potentially vulnerable Department of Defense (DoD) installations. Finally, the team would like to thank collaborators including Ashray Manepalli of the University of California-Berkeley and Adrian Albert, Alan Rhoades, and Prabhat of the Lawrence Berkeley National Laboratory for developing discussions and leading efforts to use state-of-the-art machine-learning methods for developing new ways to probe models for their sensitivity to different processes.

1.0 INTRODUCTION

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

Large infrastructure decisions are made based on average temperature and precipitation, their variability, and extreme events. Such decisions may also require many years of lead-time. General Circulation Models (GCMs), sometimes referred to as global climate models, can be a tool to assess changes in the large-scale distribution of temperature and precipitation as a function of future human activities, which are conventionally divided into discrete greenhouse gas emissions scenarios (van Vuuren et al, 2011). The range of global-scale temperature response to the forcing associated with a doubling of carbon dioxide is likely not more than 4.5 °C and has been established through a sustained focus on climate sensitivity (Charney et al, 1979; Knutti and Hegerl, 2008; Forster, 2016; Knutti et al, 2017). The response of the Earth's hydrological cycle as it adjusts to warming has also been studied, and gross features, such as the global hydrological cycle response to warming (Jeevanjee and Romps, 2018), the expansion of the troposphere (Lu et al, 2007; Hu et al, 2013), and shifts in the storm tracks (Yin, 2005; Bengtsson et al, 2006; Ulbrich et al, 2008), have been observed and have a strong theoretical basis.

These models are also blunt tools for finer-scale questions: the numerical integration of the intricate computer program that represents a GCM must balance a number of competing factors including, but not limited to, top-of-atmosphere energy balance, large-scale circulation, cloud formation and dissipation, ocean dynamics, and earth system response (Mauritsen et al, 2012; Hourdin et al, 2017; Schmidt et al, 2017). Furthermore, model development often focuses on achieving unbiased assessments of mean temperature and precipitation at larger spatial-scales and long time-scales. Because regional variability in precipitation and precipitation change across models is quite broad (Pendergrass and Hartmann, 2014), it is not known *a priori* that a given model will produce skillful temperature and precipitation projections as they relate to DoD facilities.

Nominally, model intercomparison projects serve to characterize the spread of model results associated with a given emissions scenario, but it has been recently recognized that navigating through the multi-model ensemble to understand the information contained in the ensemble's projections must consider both model errors with respect to observations and the interdependence of models (Sanderson et al, 2015; Knutti et al, 2017; Sanderson et al, 2017; Eyring et al, 2019). This navigation depends on the specific question being asked. Model democracy, where each simulation is treated as independent and unbiased, has been found to be suboptimal, primarily owing to the fact that models have not been developed independently and may even have substantial similarities in their code bases (Sanderson et al, 2015). Rather, it is essential that models being considered exhibit a level of skill with respect to regional metrics (Rupp et al, 2013). Taking California as an example, only models that reproduce the correlation and variance of mean seasonal spatial patterns, amplitude of seasonal cycle, diurnal temperature range, annual- to decadal-scale variance, long-term persistence, and Western United States regional precipitation teleconnections to El Niño Southern Oscillation are appropriate for hydroclimate studies, and only 10 of the 33 Coupled Model Intercomparison Project – Phase 5 (CMIP5) models exhibited skill

with respect to this metric (Pierce et al, 2018; Thorne et al, 2018). The 4th National Climate Assessment (USGCRP, 2017) adopted a model-weighting approach to produce, from the CMIP5 ensemble, regional-scale estimates of change in average temperature and precipitation. The results broadly indicate high confidence in predicting more warming with the Representative Concentration Pathway 8.5 (RCP8.5) scenario than the RCP4.5 scenario across the Conterminous United States (CONUS), and, with less confidence, more variability in regional precipitation for both of these scenarios.

At the same time, the Coupled Model Intercomparison Project - Phase 6 (CMIP6) is, as of this writing, in progress, with preliminary results coming online. At present, 151 modeling activities (one model running one set of experiments for one model intercomparison project constitutes an activity) have been reported to the Earth System Grid Federation (Cinquini et al, 2014) The contributing modeling centers are expected to produce between 20 and 40 Petabytes of climate model simulations as part of the CMIP6 exercise (Eyring et al, 2016). This is far too voluminous for planning purposes, and therefore, there is a strong need to curate these results to determine ranges of conditions that are likely to be faced locally. A number of different experiments are contained within CMIP6, including the Shared Socioeconomic Pathways (O'Neill et al, 2014; O'Neill et al, 2017) that have close heritage with the Representative Concentration Pathways (van Vuuren et al, 2011) of CMIP5 (Taylor et al, 2012). Other experiments including the North-American Coordinated Regional Downscaling Experiment (Giorgi et al, 2009; Jones et al, 2011; Mearns et al, 2012), the High Resolution Model Intercomparison Project (HighResMIP) (Haarsma et al, 2016), and the Vulnerability, Impacts, Adaptation, and Climate Services (VIACS) (Ruane et al, 2016) will yield simulations at the spatial resolution of interest to DoD facilities. It is expected that there will be many participating modeling centers in each of these intercomparison projects, and there is a need to determine how to interpret the range of results that arise thereof specifically in terms of DoD facility-level risk.

It is critical that the climate change signals that are being captured by state-of-the-art climate models are translated to the local level and that the process of doing so does not underestimate the risks that climate change effects may have on local infrastructure. Unfortunately, raw climate model output can exhibit biases in temperature and precipitation at regional scales, and these biases will propagate in complex and potentially nonlinear ways into subsequent models and decision-making efforts (Christensen et al, 2008; Mearns et al, 2012; Sillmann et al, 2013). To address this issue, one approach, while not without controversy (Ehret et al, 2012; Maraun 2012; Maraun 2013) is to bias-correct the raw GCM output based on historical observations (Gudmundsson et al, 2012; Teutschbein and Seibert, 2012; Teutschbein and Seibert, 2013; Pierce et al, 2015). An additional and critical step involves downscaling to address the gross mismatch in the spatial resolution on which GCMs operate and the processes that impact the local-level.

In November 2017, SERDP released a Workshop Report entitled "Nonstationary Weather Patterns and Extreme Events" and found that "[improving] modeling and tools to develop, apply, and evaluate scenarios of multiple interacting stresses, including training in underlying assumptions and application of scenarios" and "[validating] current [downscaling] methods, [developing] and [testing] new approaches, and [providing] guidance for different situations, locations, and variables" were among the greatest needs and opportunities for future research (Moss et al, 2017).

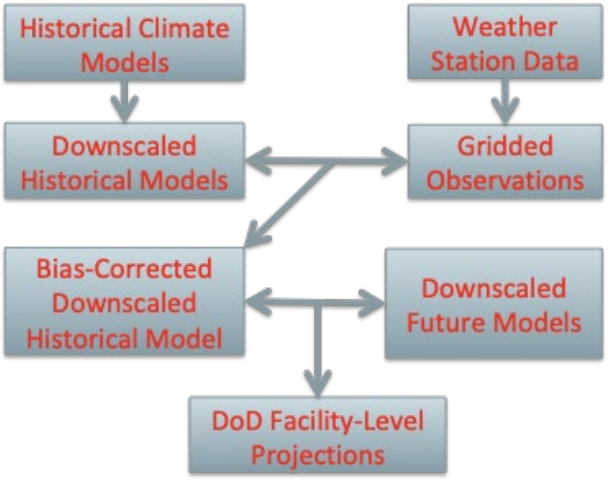

Figure 1. Flow-chart Indicating How CMIP Model Projections Are Translated into Model Facility-level Risk.

Adapted from Feldman et al, (Submitted 2019).

Figure 1 shows a flow-chart that qualitatively produces DoD facility-level projections which balance historical observations with the relative change in the downscaled, bias-corrected model between future and historical simulations.

Distributions of temperature and precipitation can be constructed based on one or more bias-corrected, downscaled models results, but each step of this process must be critically evaluated, and several questions arise:

- How should individual models and the ensemble of model results be handled to develop confidence intervals?
- Can certain model results be trusted more than others?
- When and where is it appropriate to use, in the course of developing future projections, downscaling techniques and bias-correction based on historical observations?

2.0 OBJECTIVES

This investigation, arising in response to RCSON-18-L2 "Quantifying the Effects of Environmental Nonstationarity at DoD Relevant Scales," seeks to uncover the heterogeneous performance of the GCMs with respect to nonstationarity of temperature and precipitation. The primary objective of this research is to build the foundation for a comprehensive determination of this risk by using historical observations to inform model projections of the statistical properties of temperature and precipitation, including nonstationarity, at regional scales in the CONUS.

The research is therefore directly responsive to the research objective in RCSON-18-L2 of "Discernment of relationships (e.g., mean temperature and heat wave frequency or severity) as revealed in both historical observations and climate models with an emphasis on describing model heterogeneity.

This research has been designed to lay the groundwork for tangible benefits to DoD by establishing a physical basis for clarity in (1) the risk that future changes in the distribution of temperature and precipitation that a single DoD facility may face for a set of plausible emissions scenarios and (2) the risk that multiple DoD facilities may face simultaneously in the future due to these scenarios. The research first focused on the use of traditional geostatistical techniques including spatiotemporal variography and kriging to blend historical data and GCM projections to examine the spatiotemporal behavior of temperature and precipitation over the CONUS in future emissions scenarios.

As a corollary to this primary objective, we pursued a secondary objective. As per the needs expressed in RCSON-18-L2, we explored the assumptions inherent in pursuing a statistical approach to understanding the risks associated with nonstationarity and whether physical modeling and understanding are needed to bringing clarity to interpreting these risks under future climate change.

3.0 TECHNICAL APPROACH

The null hypothesis of this investigation was that the statistics of temperature and precipitation and their covariability in CONUS on regional scales have been and will continue to be stationary. At the start of the investigation, the research team pursued a purely statistical approach to this question. That is, the research team sought to develop projections of changes in the distributions of temperature and precipitation based on a combination of model projections where the projections were weighted according to the model's historical performance relative to observations.

Pursuant to this strategy, the research team developed a multi-linear regression based on the three-dozen models from the Couple Model Intercomparison Project – Phase 5 (CMIP5) (Taylor et al 2012) archive that was downscaled using the Localized Constructed Analogues (LOCA) technique (Pierce et al, 2014). The research team developed a regression approach of the historical model simulations by comparing those against a gridded hydrometeorological product (Livneh et al, 2013). Biases were derived as a function of quantile for each model and this was used as a basis to develop a multi-linear regression. As part of this process, the research team were able to develop exceedance probabilities for the projections based on a generalized extreme value (GEV) distribution. A number of techniques were also developed to manage potential numerical instabilities associated with the construction of these projections. A flow-chart of this approach is displayed in Figure 2 and described in further detail in Feldman et al (2019, Submitted) and Tadić and Biraud (2019, Submitted). The research team also developed a set of more advanced statistical techniques to manage potential changes in the spatiotemporal correlation in temperature and precipitation, through a copula-based technique (Tadić and Biraud, 2019, Submitted).

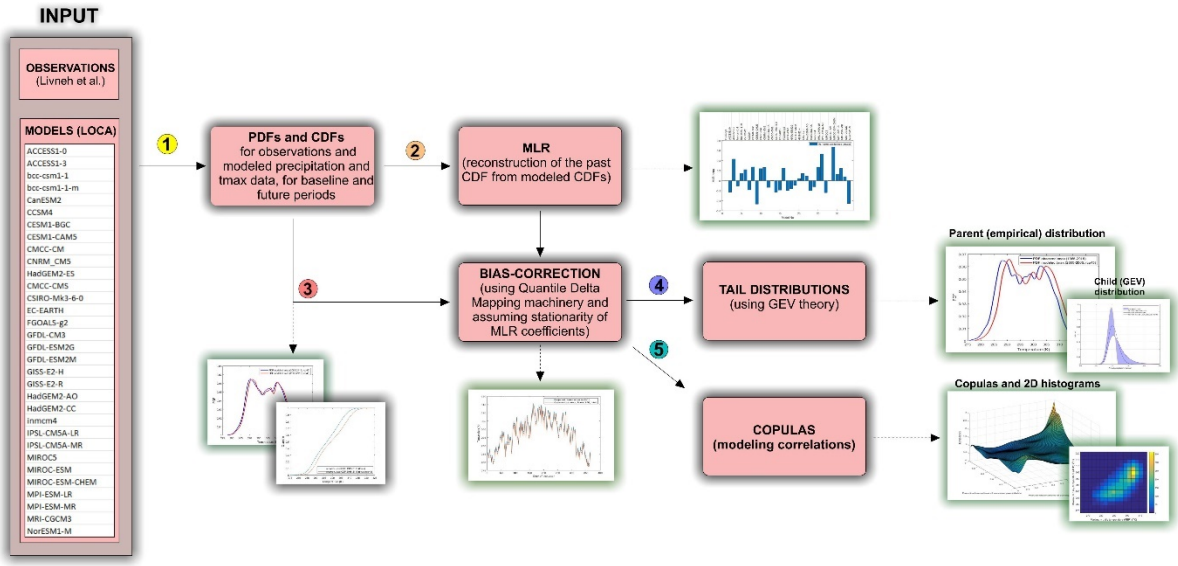

Figure 2. Flow Diagram of a Regression-based Analysis.

(1) Input data (observational and model data) are first converted into empirical Probability Distribution Functions and Cumulative Distribution Functions (CDFs). (2) Multi-linear regression (MLR) coefficients are constructed to reconstruct observational CDFs using model CDFs. (3) Quantile Delta Mapping (Cannon et al, 2015) Quantile Delta Mapping is applied to develop future CDFs. (4) We apply GEV theory and examine change in the likelihood of extreme events. Finally, (5) copula theory is applied to assess whether climate change increases risk of simultaneous or collocated extreme events. From Tadić and Biraud (2019, Submitted).

With the mechanics of what amounts to a pure statistically-based approach to developing projections to assess future changes for a given location and correlated risks for multiple locations, the research team sought to revisit key assumptions that were made in the course of constructing this statistical model. Based on key feedback at the SERDP Symposium from Dr. Linda Mearns of the National Center for Atmospheric Research (NCAR), and following the suggestions of Dr. Michael Wehner of Lawrence Berkley National Lab, the research team recognized the importance of adopting a more cautious approach to performing analysis on a downscaled multi-model ensemble. Specifically, key assumptions underlying the approach were assessed, including that models are independent and that the stationarity assumptions in downscaling techniques are tenable. Through discussions with the co-performers on this proposal, it became clear that a careful intercomparison of different downscaling solutions provides information not just on the downscaling technique, but on the parent model as well.

A deeper understanding of the strengths and weaknesses of the parent models and downscaling techniques, which can be informed by the intercomparison of different downscaling solutions, is needed, along with a robust test of stationarity assumptions in downscaling. With a multi-pronged approach to testing parent models and downscaling techniques, a rational basis for weighting or deweighting climate model projections and downscaling techniques can be developed. The research team's original approach to developing projections represents a default set of assumptions regarding the a priori equality in the independence, uniqueness, and plausibility of model projections and stationarity in statistical downscaling relationships between the 20th and 21st Century.

These assumptions can be tested and an overview for doing so is in the conclusion of the Executive Summary. A detailed approach for accomplishing this is found at the conclusion of this report.

4.0 RESULTS AND DISCUSSION

Using the approach outlined in Figure 2, the research team first focused on the development of projections of daily precipitation and surface air temperature around two selected DoD facilities: Travis Air Force Base (TAFB) near Fairfield, California, and Fort Bragg (FBR) in North Carolina. Below is a summary of average and extreme changes for these two locations. The changes of mean temperature are not necessarily monotonic with time and emissions, though they are large, and the changes in mean precipitation are moderate.

Table 1. Changes in Average Daily Precipitation and Maximum Daily Surface Temperature Values (Future-Past) for TAFB for Three Future Periods Centered Around Years 2020, 2050 and 2100, Derived from Bias-corrected Future Projections Using MLR and Past Observations, for Two Emission Scenarios (RCP4.5 and RCP8.5).

	TAFB			
	RCP45		RCP85	
Year (centered around)	Precipitation (mm/day) (delta)	T_{max} (K) (delta)	Precipitation (mm/day) (delta)	T_{max} (K) (delta)
2020	+0.1	+1.1	+0.2	+1.1
2050	-0.0	+2.0	+0.0	+1.9
2100	-0.1	+2.2	+0.0	+2.7

Table 2. Changes in Average Daily Precipitation and Maximum Daily Surface Temperature Values (Future-Past) for FBR for Three Future Periods Centered Around Years 2020, 2050 and 2100, Derived from Bias-corrected Future Projections Using MLR and Past Observations, for Two Emission Scenarios (RCP4.5 and RCP8.5).

	FBR			
	RCP45		RCP85	
Year (centered around)	Precipitation (mm/day) (delta)	T_{max} (K) (delta)	Precipitation (mm/day) (delta)	T_{max} (K) (delta)
2020	-0.5	+0.9	-0.3	+0.1
2050	-0.4	+1.2	-0.3	+1.5
2100	-0.5	+1.6	-0.4	+2.2

This approach can be used to develop estimates of changes in extreme values of precipitation and daily maximum temperature in the 21st Century and associated confidence intervals using GEV parameter estimates. These changes are a function of emissions scenario, which becomes a large driver of the divergence in estimates by the end of the 21st Century (Hawkins and Sutton, 2009).

Table 3. The Ratio of Probability of the Extreme Event in the Future that Corresponded to p<0.01 in the Past, and Its Probability for Three Future Periods (Centered around years 2020, 2050 and 2100), for Both TAFB and FBR Locations Under Two Emission Scenarios, Based on GEV Distribution.

The uncertainty is reported as 95% Confidence Intervals (CI).

	TAFB				**FBR**			
	Precipitation (mm/day) (95% CI)		**T_{max} (K) (95% CI)**		**Precipitation (mm/day) (95% CI)**		**T_{max} (K) (95% CI)**	
	RCP4.5	**RCP8.5**	**RCP4.5**	**RCP8.5**	**RCP4.5**	**RCP8.5**	**RCP4.5**	**RCP8.5**
2005-2035	1.09 (0-16.5)	2.77 (0-20)	4.35 (0-21.4)	4.63 (0-21.5)	0.30 (0-10)	0.47 (0-8)	7.58 (0-35.5)	1.86 (0-26)
2035–2065	0.20 (0-11)	1.87 (0-17)	11.35 (0-31.1)	11.63 (0-31.5)	0.23 (0-9)	0.33 (0-8)	11.83 (0-38)	17.65 (0-41.5)
2085–2115	0.06 (0-10)	2.97 (0-19)	13.60 (0-33.4)	22.45 (0.7-42.5)	0.21 (0-9)	0.25 (0-9)	19.98 (0-45)	31.04 (0-52.5)

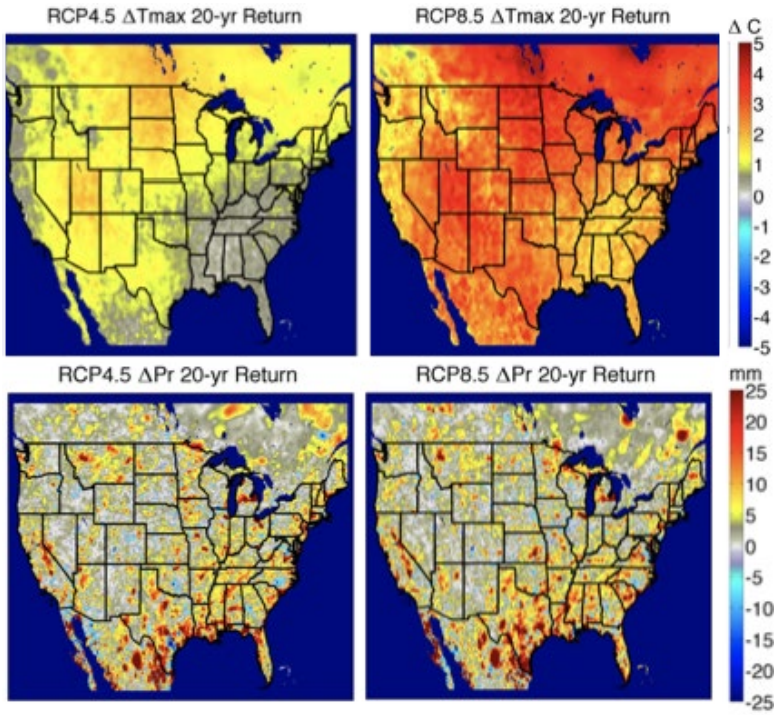

Figure 3. Change in 20-year Daily Return Value of Maximum Surface Air Temperature and Precipitation Based on the MLR Approach Across the CONUS for Two Emissions Scenarios from 1985-2015 to 2070-2100.

This analysis can be extended to develop distributions across the CONUS. We summarize these results by showing the change in the 20-year daily return value of maximum daily temperature and precipitation in Figure 3.

This approach provides projections, but the steps shown in Figure 1 need to be approached cautiously. In point of fact, seemingly straightforward tasks, such as the development of gridded historical data from a discrete set of weather station observations, can introduce biases. Co-Performers Risser and O'Brien recently published a paper showing that gridding observations will tend to underestimate extreme precipitation because that precipitation varies spatially, and averaging high and low precipitation data will dampen extremes (Risser et al, 2019a). A better approach is to use station data to estimate extreme precipitation statistics and develop spatial modeling to produce gridded observational products.

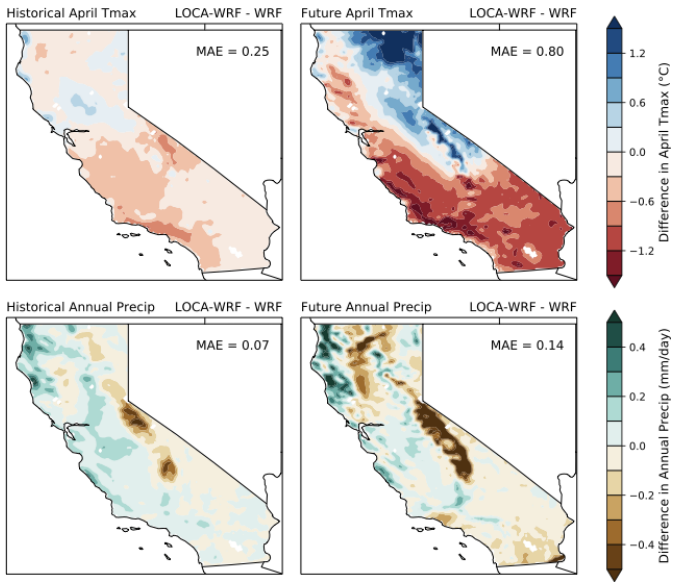

Figure 4. Differences Between Weather Research and Forecasting (WRF) and LOCA-WRF Projections for Historical and Future April Maximum Temperature (Panels a and b, Respectively) and for Historical and Future Annual Mean Precipitation (Panels c and d, Respectively).

Unit for (a) and (b) is degrees C. Unit for (c) and (d) is millimeters per day. The mean absolute error (MAE) is reported in the upper right margin of each panel. From Walton et al, JAMC, In Review.

Additionally, recent work by Walton et al, (JAMC, In Review), as shown in Figure 4, indicates that the regression approach is suboptimal. An intercomparison of different downscaling solutions across California for historical and future simulations reveals a divergence that only appears later in the 21st Century. A closer look reveals that the dynamically-downscaled solution is qualitatively and quantitatively different from the statistical downscaling solution. The reason that the historical performance of the statistical downscaling solution does not translate well into future performance requires careful analysis. Walton et al, (JAMC, In Review) conducted this analysis and found that, for California, the downscaling methodology itself is only a symptom of a larger problem. The primary driver for the divergence in downscaling solutions in future projections is that the parent model does not capture changes in snow cover well, which greatly impact temperatures and does not capture changes in North American Monsoon dynamics, which affects Southeastern California.

These findings show that the blind development of regression solutions based on historical biases is suboptimal, because the historical biases of some models are not directly translatable to future biases.

Rather, it is the process representations of snow albedo and monsoons that must be considered in the development of projections. It is precisely this nonstationarity that RCSON- 18-L2 seeks to address head on. The suboptimal performance of MLR-based projections, where the regression is based on the univariate performance of the models in the historical record, is only revealed with the intercomparison of dynamical and statistical downscaling solution.

5.0 IMPLICATIONS FOR FUTURE RESEARCH AND BENEFITS

This research explored the landscape of developing projections of surface air temperature and precipitation at the DoD-facility level from an ensemble of downscaled climate model projections. The research demonstrated that this landscape is not easily navigated because climate models have representation errors in the processes that can lead to nonstationarity and statistical downscaling solutions are constructed explicitly with a stationarity assumption that has not been tested.

The primary benefit to DoD of this seed project's research is to detail the scale of the work that is needed to address in a comprehensive fashion specific research tasks that go beyond a purely statistical approach to developing climate model projections from an ensemble of model results to provide insight into the models with which DoD should have increased or diminished confidence. The challenges of using downscaled projections for infrastructure and operations planning should not be underestimated, because they are generally completely opaque to the end-user. These end-users can easily have either too much or too little confidence in the model projections. For the former case, the provision of a set of numbers, and even error bars, that can be fed into end-user models is a straightforward exercise, but the danger here is that those numbers can be misleading if they do not consider the processes that break stationarity.

For the latter case, the esoteric nature of climate models and downscaling can preclude their adoption even though it is very likely that long-term shifts in surface air temperature and precipitation distributions will occur. In this situation, planning exercises either use historical data that do not include nonstationarity or generic information on increased variability as a placeholder for a deeper understanding of the risks associated with climate change which may or may not be relevant to a specific location.

Fortunately, there are tractable solutions to these challenges. Focused research efforts which develop a deeper understanding of the processes that "break" stationarity and evaluate model projections accordingly are needed. Additionally, it is critical that the researchers engaged in developing climate model projections bridge the communication divide between the modeling community and end-users so that they have the appropriate level of confidence in these projections. Ultimately, end-users need to understand the challenges and opportunities with using climate model projections for infrastructure and operations planning and so that they can utilize model projections judiciously. It is incumbent on those with knowledge and expertise in climate model projections to guide the utilization of these data for end-users.

6.0 LITERATURE CITED

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

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

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher (2008), On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, **35**, L20709, doi:10.1029/2008GL035694.

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

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert, (2012) HESS opinions: ''Should we apply bias correction to global and regional climate model data?'' *Hydrol. Earth Syst. Sci.*, **16**, 3391–3404, doi:10.5194/hess-16-3391-2012.

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

Eyring, V., et al., (2019) Taking climate model evaluation to the next level, *Nature Climate Change*, **9**, 102-110, doi:10.1038/s41558-018-0355-y.

Feldman, D.R., J.M. Tadić, W. Arnold, A.Schwarz (2019) Establishing a Range of Extreme Precipitation Estimates in California for Planning in the Face of Climate Change, *Water Resources Research* (Submitted).

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

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

Gudmundsson, L., J. Bremnes, J. Haugen, and T. Engen-Skaugen, (2012) Technical note: Downscaling RCM precipitation to the station scale using statistical transformations comparison of methods. *Hydrol. Earth Syst. Sci.*, **16,** 3383–3390, doi:10.5194/ hess-16-3383- 2012.

Haarsma, R.J., et al., (2016) High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, *Geosci. Model Dev.*, **9**, 4185-4208, doi:10.5194/gmd-9-4185-2016.

Hourdin, F., T. Mauritsen, A. Gettelman, J. Golaz, V. Balaji, Q. Duan, D. Folini, D. Ji, D. Klocke, Y. Qian, F. Rauser, C. Rio, L. Tomassini, M. Watanabe, and D. Williamson, (2017) The Art and Science of Climate Model Tuning. *Bull. Amer. Meteor. Soc.*, **98**, 589–602, doi:10.1175/BAMS-D-15-00135.1.

Jeevanjee, N. & D.M. Romps (2018) Mean precipitation change from a deepening troposphere, *Proc. Nat. Acad. Sci. USA*, **115**, 11465-11470, doi:10.1073/pnas.1720683115.

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

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

Knutti, R., M. A. A. Rugenstein, G.C. Hegerl, (2017), Beyond equilibrium climate sensitivity, *Nature Geoscience,* **10**, 727-736, doi:10.1038/ngeo3017.

Livneh, B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer, and D.P. Lettenmaier (2013) A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *J. Climate,* **26**, 9384– 9392, doi:10.1175/JCLI-D-12-00508.1.

Lu, J., G. A. Vecchi, and T. Reichler (2007), Expansion of the Hadley cell under global warming, *Geophys. Res. Lett.*, 34, L06805, doi:10.1029/2006GL028443.

Maraun, D. (2012) Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophys. Res. Lett.*, **39**, L06706, doi:10.1029/2012GL051210.

Maraun, D., (2013) Bias correction, quantile mapping, and downscaling: Re-visiting the inflation issue. *J. Climate*, **26**, 2137–2143, doi:10.1175/ JCLI-D-12-00821.1.

Mauritsen, T., et al. (2012). Tuning the climate of a global model, *Journal of Advances in Modeling Earth Systems*, **4**, doi:10.1029/2012MS000154.

Milly, P.C.D., J. Betancourt, M. Falkenmark, R.M. Hirsch, W.Z. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer. 2008. Stationarity Is Dead: Whither Water Management? *Science*. **319**, 573–574, doi:10.1126/science.1151915.

Mearns, L. O., et al (2012) The North American Regional Climate Change Assessment Program: Overview of phase I results. *Bull. Amer. Meteor. Soc.*, **93**, 1337–1362, doi:10.1175/ BAMS-D- 11-00223.1.

Moss, R.H., et al, (2017) Nonstationary Weather Patterns and Extreme Events: Informing Design and Planning for Long-Lived Infrastructure, ESTCP Project RC-201591, https://serdp-estcp.org/content/download/46665/436260/file/Nonstationarity%20RC%20Workshop%20Report.pdf

O'Neill, B.C., Kriegler, E., Riahi, K. et al. (2014) A new scenario framework for climate change research: the concept of shared socioeconomic pathways, *Climatic Change*, **122**, 387, doi:10.1007/s10584-013-0905-2.

O'Neill, B.C., et al. (2017) The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century, *Global Environmental Change*, 42, 169-180, doi: 10.1016/j.gloenvcha.2015.01.004.

Oyler, J.W., S.Z. Dobrowski, Z.A. Holden, and S.W. Running, (2016) Remotely Sensed Land Skin Temperature as a Spatial Predictor of Air Temperature across the Conterminous United States. *J. Appl. Meteor. Climatol.,* **55**, 1441–1457, doi:10.1175/JAMC-D-15-0276.1.

Pendergrass, A.G. and D.L. Hartmann (2014) Changes in the Distribution of Rain Frequency and Intensity in Response to Global Warming. *J. Climate,* **27**, 8372–8383, doi:10.1175/JCLI-D-14- 00183.1.

Pierce, D. W., T. Das, D. R. Cayan, E. P. Maurer, N. L. Miller, Y. Bao, M. Kanamitsu, et al. (2012) Probabilistic Estimates of Future Changes in California Temperature and Precipitation Using Statistical and Dynamical Downscaling. *Climate Dynamics*, **40** (3-4): 839–56, doi: 10.1007/s00382-012-1337-9.

Pierce, D. W., D. R. Cayan, and B. L. Thrasher, (2014) Statistical downscaling using Localized Constructed Analogs (LOCA). *J. Hydromet.*, **15**, 2558-2585, doi:10.1175/JHM-D-14-0082.1.

Pierce, D. W., D. R. Cayan, E. P. Maurer, J. T. Abatzoglou, and K. C. Hegewisch, (2015) Improved bias correction techniques for hydrological simulations of climate change. *J. Hydrometeorology*, 16, 2421-2442. doi:10.1175/JHM-D-14-0236.1.

Pierce, D. W., D. R. Cayan, J. F. Kalansky. (Scripps Institution of Oceanography). 2018. Climate, Drought, and Sea Level Rise Scenarios for the Fourth California Climate Assessment. California's Fourth Climate Change Assessment, California Energy Commission. Publication number: CCCA4-CEC-2018-006.

Risser, M.D., Paciorek, C.J., O'Brien, T.A., Wehner, M.F., Collins, W.D., (2019a) A Probabilistic Gridded Product for Daily Precipitation Extremes Over the United States, *Climate Dynamics*, doi: 10.1007/s00382-019-04636-0.

Risser, M.D., Paciorek, C., O'Brien, T.A., Wehner, M.F., Collins, W.D. (2019b) Detected changes in precipitation extremes at their native scales derived from in situ measurements, *J. Climate*, (Accepted) https://arxiv.org/abs/1902.05977.

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

Rupp, D. E., J. T. Abatzoglou, K. C. Hegewisch, and P. W. Mote (2013), Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA, *J. Geophys. Res. Atmos.*, **118**, 10,884–10,906, doi:10.1002/jgrd.50843.

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

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

Schmidt, G.A., D. Bader, L.J. Donner, G.S. Elsaesser, J.-C., Golaz, C. Hannay, A. Molod, R.B. Neale, S. Saha (2017) Practice and philosophy of climate model tuning across six US modeling centers, *Geosci. Mod. Dev.*, 10, 3207-3223, doi:10.5194/gmd-10-3207-2017.

Sillmann, J., V. Kharin, X. Zhang, F. Zwiers, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.*, **118**, 1716–1733, doi:10.1002/jgrd.50203.

Tadić, J.M. and S.C. Biraud (2019) Towards More Accurate Modeling of Future Weather: Challenges of Downscaling, Bias Correction, Correlation, Risk Assessment, and Non-Stationarity, *Journal of Applied Meteorology and Climatology*, (Submitted).

Taylor, K.E., R.J. Stouffer, and G.A. Meehl, (2012) An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.,* **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Teutschbein, C., and J. Seibert, (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.*, **456–457**, 12–29, doi:10.1016/j.jhydrol.2012.05.052.

Teutschbein, C. and Seibert, J. (2013) Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrol. Earth Syst. Sci.*, **17**, 5061-5077, doi:10.5194/hess-17-5061-2013.

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

Timmermans, B., Wehner, M., Cooley, D., O'Brien, T, Thibaud, E., & Krishnan, H. (2018). An evaluation of the consistency of extremes in gridded precipitation data sets. *Climate Dynamics*, 1-20, doi:10.1007/s00382-018-4537-0.

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

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

Wang, M., and P. Ullrich. (2018) Marine Air Penetration in California's Central Valley: Meteorological Drivers and the Impact of Climate Change. *Journal of Applied Meteorology and Climatology*, **57**(1), 137–54. doi:10.1175/JAMC-D-17-0089.1.

Yin, J. H. (2005), A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophys. Res. Lett.*, **32**, L18701, doi:10.1029/2005GL023684.