Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections

Release of Downscaled LOCA CMIP5 Hydrology

Bureau of Reclamation

California Energy Commission

Climate Analytics Group Climate Central Lawrence

Livermore National Laboratory NASA's Ames Research Center, Moffett Field, CA

Santa Clara University

Scripps Institution of Oceanography

U.S. Army Corps of Engineers

U.S. Geological Survey

Cooperative Institute for Research in Environmental Sciences

Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections

Release of Downscaled LOCA CMIP5 Hydrology

June 2020

Authors:

Julie Vano, Joseph Hamman, Ethan Gutmann, Andrew Wood, Naoki Mizukami, Research Applications Laboratory, National Center for Atmospheric Research

Martyn Clark, University of Saskatchewan

David Pierce, Daniel Cayan, U.S. Geological Survey and University of California- San Diego/Scripps Institution of Oceanography

Cameron Wobus, Lynker Technologies

Kenneth Nowak, Bureau of Reclamation

Jeffrey Arnold, U.S. Army Corps of Engineers

Prepared for:

Users of the downscaled CMIP5 Climate and Hydrology Projections available at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections .

Table of Contents

Figures and Tables

Abbreviations and Acronyms

BCSD – Bias-Correction and Spatial Disaggregation

BCCA – Bias-Correction Constructed Analogs

CMIP – Climate Model Intercomparison Project

CONUS – COnterminous United States

ET – Evapotranspiration

GCM – Global Climate Model (or General Circulation Model)

GDO – Green Data Oasis

LOCA – LOcalized Constructed Analogs

RCP – Representative Concentration Pathways

SWE – Snow Water Equivalent

VIC – Variable Infiltration Capacity hydrologic model

Citations of these Projections:

When publishing research based on projections from this archive, please include two acknowledgements:

1. Acknowledge the superseding effort:

   a. For Coupled Model Intercomparison Project phase 3 (CMIP3), the following is language suggested by the CMIP3 archive hosts at the Program for Climate Model Diagnosis and Intercomparison (PCMDI):

   "We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and the World Climate Research Programme (WCRP) Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy."

   PCMDI also requests that in first making reference to the projections from this archive, please first reference the CMIP3 dataset by including the phrase "the World Climate

Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset." Subsequent references within the same publication might refer to the CMIP3 data with terms such as "CMIP3 data," "the CMIP3 multi-model dataset," "the CMIP3 archive," or the "CMIP3 dataset."

b. For Coupled Model Intercomparison Project phase 5 (CMIP5), the model output should be referred to as "the CMIP5 multi-model ensemble [archive/output/results/of simulations/dataset/ ...]." In publications, you should include a table (listing the models and institutions that provided model output used in your study. In this table, and as appropriate in figure legends, you should use the CMIP5 "official" model names found in "CMIP5 Modeling Groups and their Terms of Use" (https://pcmdi.llnl.gov/mips/cmip5/docs/CMIP5_modeling_groups.pdf). In addition, an acknowledgment similar to the following should be included in your publication:

"We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table XX of this paper) for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals."
where "Table XX" of your paper should list the models and modeling groups that provided the data you used. In addition, it may be appropriate to cite one or more of the CMIP5 experiment design articles listed on the CMIP5 reference page.

2. Second, generally acknowledge this archive as "Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections" archive at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections . To reference specific statistically downscaled climate and hydrology data from the archive, please use the following references:

a. *for BCSD CMIP3 climate*: Maurer, E.P., L. Brekke, T. Pruitt, and P.B. Duffy, 2007. Fine-resolution climate projections enhance regional climate change impact studies, Eos Trans. AGU, 88(47), 504, doi:10.1029/2007EO470006.

b. *for BCSD CMIP3 hydrology*: Reclamation, 2011. West-Wide Climate Risk Assessments: Bias-Corrected and Spatially Downscaled Surface Water Projections, Technical Memorandum No. 86-68210-2011-01, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 138 p, https://www.usbr.gov/watersmart/docs/west-wide-climate-risk-assessments.pdf.

c. *for BCSD CMIP5 climate*: Provide citation to: Reclamation, 2013. Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs. U.S. Department of the Interior, Bureau of Reclamation, 104 p., available at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_climate.pdf .

d. *for BCSD CMIP5 hydrology*: Reclamation, 2014. Downscaled CMIP3 and CMIP5 Hydrology Projections - Release of Hydrology Projections, Comparison with Preceding Information, and Summary of User Needs. U.S. Department of the Interior, Bureau of

Reclamation,104 p., available at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_ climate.pdf .

e. *for the locally constructed analogs method (LOCA) CMIP5 projections*: Pierce, D. W., D. R. Cayan, and B. L. Thrasher, 2014. Statistical Downscaling Using Localized Constructed Analogs (LOCA)*, Journal of Hydrometeorology, 15(6), 2558-2585, doi:10.1175/JHM-D-14-0082.1; and 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, v. 16, p. 2421-2442. doi:10.1175/JHM-D-14-0236.1

f. *for LOCA CMIP5 hydrology*: Vano, J., J. Hamman, E. Gutmann, A. Wood, N. Mizukami, M. Clark, D. Pierce, D. Cayan, C. Wobus, K. Nowak, and J. Arnold. (June 2020). Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections - Release of Downscaled LOCA CMIP5 Hydrology. 96 p, available at: https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/LOCA_BCSD_hydrology_tech_memo.pdf .

Executive Summary

This memorandum describes and compares climatology and hydrology outputs created with the LOcalized Constructed Analogs (LOCA) method of empirical-statistical downscaling [Pierce et al., 2014] and the Bias Correction and Spatial Disaggregation (BCSD) [Wood et al., 2004] method. The descriptions and comparisons are made from downscaled climatology generated for the Conterminous United States (CONUS) and hydrology generated using climatology from LOCA and BCSD to drive the Variable Infiltration Capacity (VIC) hydrologic model [Liang et al., 1996, Gao et al., 2010]. Hydrologic variables are the chief focus here: evapotranspiration (ET), runoff (surface + baseflow), snow water equivalent (SWE), along with the meteorological variables temperature and precipitation both for individual projections and ensemble averages. Modeled output is presented as long-term 30-year annual averages (section 3.1), seasonal changes (3.2), and daily statistics (3.3).

The objective of this report is to summarize the overall, not exhaustively specific, performance of the two methods for select climatological and hydrologic variables as an aid to researchers and practitioners who will use the modeled outputs. We have not diagnosed and explained all model behavior seen across the methods nor ranked the methods based on these performance comparisons. Researchers and practitioners are reminded again to make their selection of downscaling method(s) based on all of their requirements.

We found both LOCA- and BCSD-based datasets realistically reproduce the mean climate state of the observational dataset they are trained on. In most historical comparisons, these two methods provide similar results though we highlight locations and variables where differences arise (Figure 1). In particular, differences, especially in hydrologic variables, can be frequently attributed to the observational dataset used to train the downscaling method. For example, these methods differ in the number of days with precipitation, which can substantially affect ET. Differences in modelled runoff between LOCA and BCSD in the historical period (1970-1999) can be substantial but they are consistent with the differences between modelled runoff using their respective training observation datasets.

For the future projections, the two methods generally provide similar patterns of change. Here the future period is defined as (2070-2099). Differences across hydrologic variables from the methods vary with location and variable of interest, and can be substantial. LOCA generally has less change in ET and runoff (Figure 2,3), and relatively less runoff in the future, likely resulting from the fact that LOCA has less change in average precipitation compared to BCSD, particularly in mountain regions and the Southeast. This difference is largely due to the bias correction the two methods use. Bias correction approaches can modify the climate change signal (i.e., relative difference between historical and future simulations) as a side-effect. LOCA intentionally attempts to preserve the original global climate model (GCM)-projected changes in precipitation during bias correction, while BCSD applies a bias correction based on historical values alone, which shifts the multi-decadal future change based on how GCM variability compares to the observed historical variability. Notably, LOCA explicitly projects changes in daily statistics, such as the number of wet days, based on GCM projected changes, and these can alter the changes in ET and runoff.

Future projections with both methods indicate more extreme precipitation and runoff and more frequent extreme events. The development of LOCA was driven in part by a desire to better represent changes in extreme precipitation compared to an earlier constructed analog method, BCCA [Maurer et al., 2010]. We find that there is less variability across GCMs in LOCA's change in extreme precipitation magnitudes relative to BCSD; LOCA generally projects more of

an increase in extreme precipitation magnitudes outside of mountain regions than BCSD, but, like BCSD, it also projects a step-change between historical and future periods in the frequency of extreme events, which is likely to be a statistical artifact. For some daily statistics, e.g., wet day fraction, LOCA predicts more changes than BCSD because BCSD does not explicitly predict changes at the daily time scale.

Comparisons for future conditions include ones for RCP 4.5 and RCP 8.5, confirming again that modeled differences in precipitation across downscaling methods are generally smaller than differences across the two emissions scenarios; differences in temperature are even smaller than for precipitation. Similarly, direct comparisons across individual GCMs reveal that the differences between GCMs are often much larger than the differences between LOCA and BCSD downscaling methods, particularly for changes in mean precipitation.

It is important to note that both methods are direct statistical downscaling of GCM precipitation and temperature. While the differences between them may be relatively small, if a GCM does not represent important aspects of the regional climate, the downscaled projected changes may not be reliable either. Users are encouraged to explore the physical meaning behind any changes projected to understand the reliability of those projections. Further, each downscaled GCM projection (regardless of GCM or downscaling method) should be considered a possible future not a specific prediction.

This report adds to existing information provided in "Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs" [Reclamation, 2013] and its Addendum "Release of Downscaled CMIP5 Climate Projections (LOCA) and Comparison with Preceding Information" [Reclamation, 2016]. For more complete descriptions of the Climate Model Intercomparison Project (CMIP) archives and LOCA and BCSD methods, please refer to these earlier reports and original method citations listed below in Section 2 About the Downscaled Projections.

1. Introduction

The purpose of this memorandum is to present an overview of the representation of hydrology provided in a new set of hydrologic projections derived from the LOCA downscaling method and the VIC hydrologic model. This memorandum pays particular attention to the differences between these new hydrologic projections and previous projections based on the BCSD downscaling method. The differences between the downscaled datasets are briefly outlined so those using the information can better understand how the downscaling method selected might affect their application.

The hydrologic projections of future climate provided at the Green Data Oasis (GDO) site [gdo-dcp.ucllnl.org] are derived with a chain of models that links global climate models (GCMs) to regional downscaling methods to hydrologic models. GCMs simulate climate change at typical spatial resolutions of 100 km or more, which is too coarse to capture local changes in meteorology and hydrology. To provide climate change information that represents local-scale features of importance to water resources, GCM output is often downscaled to a finer spatial resolution. Downscaling can be done in multiple ways (see overviews provided by *Wilby et al.* [2004]; *Fowler et al.* [2007]; *Gutmann et al.* [2014]; *Mearns et al.* [2014]) that include statistical methods, which rely on observed statistical relationships between coarse- and fine-spatial resolution datasets, and dynamical methods, which use regional climate models with global climate model boundary conditions to represent process-based physical relations between scales. Both approaches are useful. Statistical methods are computationally efficient but are trained on historical data, and so might have trouble simulating conditions far outside historical norms. Dynamical methods attempt to simulate more physics and processes and so may do a better job simulating conditions different than historically seen, but require considerably more computational resources, which limits their feasibility in creating large-ensemble, high-resolution downscaled climatology datasets. Dynamical methods are also not free of assumptions regarding how historical conditions will be carried into the future, since important physical parameterizations (for items such as cloud behavior or the simulation of the boundary layer) are usually selected based on comparison with historical data in the region of interest, and then assumed to be the most realistic parameterization for future conditions.

In this memorandum we focus on two widely used statistical methods: LOcalized Constructed Analogs (LOCA) [*Pierce et al.*, 2014] and the Bias-Correction and Spatial Disaggregation (BCSD) [*Wood et al.,* 2004], developed with a range of support from U.S. federal and state agencies. Downscaled projections using both the LOCA and BCSD methodologies have been produced over the Conterminous United States (CONUS) and are now archived at https://gdo-dcp.ucllnl.org. The domain actually extends slightly beyond CONUS to cover the Canadian portions of the Columbia and Milk river basins. Throughout this memorandum, we focus on this region of overlap, but it is important to note that the LOCA dataset actually extends to cover more of Canada and Mexico too.

While these two methods represent two relatively similar approaches in a diverse landscape of downscaling methods, they are not identical. As such, it is important to recognize what differences exist and understand the impact such differences might have on hydrologic projections.

In this memorandum, we evaluate the representation of current climate in both the LOCA and BCSD datasets, as well as the representation of projected changes in these datasets. All statistical methods are trained on a historical dataset. LOCA was trained on *Livneh et al.* [2015] (referred to as "Livneh" throughout), while BCSD was trained on *Maurer et al.* [2002] (referred to

as "Maurer" throughout). Differences between Livneh and Maurer are also noted along with their effect on the downscaled output and subsequent VIC hydrology simulations. In Section 3.1, we evaluate long-term annual averages of temperature, precipitation, evapotranspiration (ET), runoff (surface + baseflow), and snow water equivalent (SWE) representations. We first present the historical differences between LOCA and BCSD alongside the differences between their training observation datasets (Maurer and Livneh) for context. Then, we present changes in these hydroclimate variables in the projected future climate averaged across GCMs and from a sub-sample of individual GCMs. In section 3.2, we then focus on basin specific seasonal cycles of two-digit Hydrologic Unit Code (HUC2) aggregated values. In Section 3.3, we evaluate daily statistics for precipitation and runoff, including the flow duration curve for streamflow as routed through a channel network to streamflow locations presented in an earlier report [*Reclamation*, 2014].

2. About the Downscaled Projections

The LOCA and BCSD downscaling methods are described in detail in the earlier texts referenced below. We provide only a short explanation here.

2.1 LOCA description

The LOCA dataset is an important resource for the water resource management community. LOCA provides the primary source of downscaled data in the Fourth National Climate Assessment [*USGCRP*, 2018] and Fourth California State Climate Assessment [*Pierce et al.,* 2018]. It represents a substantial increase in spatial resolution (from $1/8^{th}$ to $1/16^{th}$ degree grid spacing, from about 12 km to 6 km) and is designed to improve spatial and temporal aspects of previous analog methods. In particular, it uses local spatial patterns, and has a bias correction method that includes a frequency-dependent correction to improve the simulation of natural climate variability and attempts to preserve the original GCM-predicted change in the bias-corrected result [*Pierce et al.,* 2015]. Spatial analog approaches translate regional patterns to local-scale features through regional pattern matching – where they identify the best-matching historical days (e.g., 30 days) as compared to coarsened observations and use these to describe the local-scale spatial relationship for each grid point being downscaled. In addition, LOCA selects the single best analog based on the nearest grid cells from a pool of analogs selected based on the larger regional pattern. In this LOCA dataset, the observation dataset used to develop these relationships is *Livneh et al.* [2015], and the training period used is 1950-2005. To provide hydrologic variables (including ET, runoff, SWE, among others), the downscaled climate information is used to run the VIC hydrologic model, version 4.2.c, at $1/16^{th}$ degree grid spacing. The VIC hydrologic model configuration used was developed in *Livneh et al.* [2013], which used model parameters calibrated over many years on a $1/8^{th}$ degree grid and moved them to a $1/16^{th}$ degree grid with a few modifications. The LOCA approach and dataset are described in more detail in:

- *Pierce et al.* [2014; 2015] – research articles that describe the method
- *Reclamation* [2016] - Reclamation report addendum that includes the LOCA release notes and comparison with preceding information
- http://loca.ucsd.edu

2.2 BCSD description

The BCSD dataset is the foundation of many previous studies. It was first released in 2007 (BCSD CMIP3) and was updated in 2013 (BCSD CMIP5), see references below. Similar to LOCA it uses statistical relationships to connect coarse model output to local scale observations where relationships are built using the distribution of observations at each individual grid point being downscaled. BCSD bias corrects the GCM data using monthly quantile mapping [*Panofsky and Brier*, 1968] derived from the historical period for each grid point. BCSD then selects a month from the historical record to use for the daily weather sequences and rescales them to match the quantile-mapped monthly total. Rescaling is multiplicative for precipitation and additive for temperature. In this BCSD dataset, the observation dataset used to develop these relationships is *Maurer et al.* [2002], which has a grid spacing of $1/8^{th}$ degree, about 12 km, and the training period used is 1950-1999. To provide hydrologic variables, the downscaled climate information is used to run the VIC hydrologic model, though these earlier datasets use an earlier VIC version 4.1.2 at $1/8^{th}$ degree grid spacing. While the version of VIC used in the BCSD and LOCA projections differs, tests of the different versions in the upper Colorado River

basin revealed differences in their output to be negligible when identical forcing and parameters are used. The approach and dataset are described in more detail in:

- *Wood et al.* [2004] – research article that describes the method
- *Maurer et al.* [2007] – release notes for BCSD CMIP3 climate
- *Reclamation* [2011] – report on BCSD CMIP3 hydrology
- *Reclamation* [2013] – report on BCSD CMIP5 climate
- *Reclamation* [2014] – report on BCSD CMIP5 hydrology

2.3 Summary of relevant LOCA/BCSD methodology differences

- **Timescales:** BCSD, as implemented here, is intrinsically a monthly approach, using monthly GCM data and historical analog months to produce daily weather. LOCA is intrinsically a daily approach, downscaling each GCM model day using historical analog days. BCSD can therefore produce downscaled daily data even if the original GCMs only provided monthly data (as was common in the older CMIP3 dataset, released ca. 2010) and is less influenced by systematic GCM errors in the simulated time sequence of daily weather. However, it does not simulate changes on a daily timescale, such as the projected changes in the number of wet days and mean precipitation on wet days seen in most GCMs along the West Coast of the CONUS. LOCA directly simulates daily changes predicted by the GCM, but will also inherit any systematic errors in the sequence of daily weather simulated by the GCM.
- **Analog matching:** BCSD uses domain-wide analog months, picked randomly from the historical observations and then scaled at each grid point so that the monthly mean matches the spatially interpolated monthly value from the GCM. LOCA uses a multi-scale matching approach, where the selected historical analog is the best match to the model value in both the synoptic-scale wider region around the point being downscaled and in a 1°x 1° box around the point being downscaled. BCSD will therefore tend to preserve observed spatial relationships across the domain, but will not change those spatial relationships in response to GCM projections. LOCA preserves the GCM's spatial relationships across the domain, and can track future changes in those relationships if the GCM dictates such changes.
- **Multi-variate Bias Correction:** In BCSD, monthly temperature and precipitation are bias corrected independently. In LOCA, daily temperature is bias corrected separately depending on the presence or absence of precipitation. This is done to better simulate the temperature at which precipitation occurs, which is important for simulating snowpack. The joint temperature/precipitation bias correction is less important for a monthly approach such as BCSD, since the observed historical months used as analogs in BCSD already have daily temperature/precipitation relations that are consistent with observations (by definition).
- **Bias Correction Method:** BCSD and LOCA have different approaches to bias correction, which affects the results. BCSD uses quantile mapping on the monthly values, which, as has been noted in a number of previous studies, can alter the original GCM projections of precipitation change (e.g., *Maraun,* 2013; *Maurer and Pierce,* 2014; *Cannon et al.,* 2015; *Pierce et al., 2015; Switanek et al.*, 2017). LOCA uses PresRat ("preserves ratio"), which, as the name suggests, attempts to preserve the original GCM-predicted precipitation change, and is one of several newer bias correction techniques that attempt to avoid having short-timescale variability affect the multi-decade climate change signal (e.g., *Li et al.,* 2010; *Michelangeli et al.,* 2009; *Haerter et al.*, 2011).

- **Domain:** In these specific implementations, BCSD and LOCA cover different domains with different spatial resolutions. The BCSD dataset is limited to CONUS, with extensions to cover the Canadian portion of the Columbia and Milk river basins, while the LOCA dataset (like the Livneh dataset) covers large portions of southern Canada and Northern Mexico as well. BCSD is on a 1/8th degree grid, while LOCA was created on a 1/16th degree grid.

2.4 Comparison overview

This memorandum is focused on comparisons between the two hydrology datasets; for that reason, only global climate models (GCMs) available for both (Table 1) are used. This includes 23 GCMs that have a single run r1i1p1 for each GCM for CMIP5 projections. Five models in the LOCA data archives are excluded because they do not have BCSD data available for all comparisons. Four models, CCSM4, GISS-E2-R, CESM1-CAM5, and FGOALS-g2, are also excluded as they used different historical period simulations in the BCSD and LOCA products (described in section 3.4). As a result, the comparisons between LOCA and BCSD are not influenced by differences in the driving GCMs, and there are enough other GCMs that these exclusions are not expected to substantially influence the ensemble average. For both meteorology and hydrology we refer to the Maurer and Livneh meteorology datasets and to the VIC hydrologic simulations forced with them as "observed". We refer to meteorology and VIC hydrologic simulations from the GCM historical period as "modeled historical".

Table 1 Downscaled model simulations compared.

ACCESS1-0	HadGEM2-CC
bcc-csm1-1	HadGEM2-ES
bcc-csm1-1-m	inmcm4
CanESM2	IPSL-CM5A-MR
CESM1-BGC	MIROC-ESM
CMCC-CM	MIROC-ESM-CHEM
CNRM-CM5	MIROC5
CSIRO-Mk3-6-0	MPI-ESM-LR
GFDL-CM3	MPI-ESM-MR
GFDL-ESM2G	MRI-CGCM3
GFDL-ESM2M	NorESM1-M
HadGEM2-AO	

Details of the home institutions that provided output used in this study for each of the listed models can be found at https://pcmdi.llnl.gov/mips/cmip5/docs/CMIP5_modeling_groups.pdf.

3. Comparisons of LOCA and BCSD

This section compares and contrasts the LOCA and BCSD datasets, focusing on large scale hydro-climatological features; other differences may be apparent at finer scales. As seen in the figures below, in general, the patterns are very similar (such that spatial averages often appear identical). When, however, values are investigated for particular locations or more specific time periods, differences are larger. Figures below are intended to display examples. Underlying data is available at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections for further exploration. Plotting and analysis scripts used here are available at https://github.com/NCAR/LOCA_Downscaling_Analysis.

3.1 Long-term 30-year annual averages

Annual averages are shown across the CONUS and the time periods are defined for the historical as 1970 to 1999 and the future as 2070 to 2099. For comparison purposes, LOCA datasets are aggregated to the same 1/8th degree grid as BCSD datasets[1]. Figures below demonstrate downscaling output for averages from the 23-ensemble members and for a subset of individual GCMs. This subset is included as examples throughout, but similar plots could be drawn for each of the 23.

3.1.1. Synthesis

This memorandum includes a variety of figures to compare and contrast the performance of LOCA and BCSD downscaling methods for different hydroclimate variables, using output from a range of GCMs and two emission scenarios. We begin by showing three synthesis figures that include only ensemble averages, the composition of which is described in subsequent sections. While such a broad synthesis is valuable, it is also important to recognize that averages are composed of individual models, particularly because hydrologic evaluations that use these downscaled datasets are a collection of individual model simulations, which are then evaluated as an ensemble (hydrology is not simulated using the ensemble average of temperature and precipitation).

The synthesis figures show historical (Figure 1) and future changes (Figure 2, 3) for five hydroclimate variables. Each figure has the 23-member ensemble of modeled GCM output downscaled using BCSD on the left column, downscaled using LOCA in the middle, and the difference between LOCA and BCSD on the right. This structure is maintained in subsequent figures throughout. We usually display differences as total magnitudes, instead of percent changes, but have plotted percent changes in some cases to convey relative differences as well (see Appendix A).

Figure 1 provides a synthesis of the historical ensemble averages across meteorological (temperature and precipitation) and hydrologic (ET, runoff, and SWE) variables for each downscaling method and their differences. Hydrologic variables of ET, runoff, and SWE are output from different implementations of the VIC hydrologic model (e.g., different resolutions and historical datasets). These plots have similar spatial patterns to the observed and modeled historical simulations (described more in section 3.1.2 and 3.1.3). This side-by-side comparison shows that the LOCA (and Livneh) VIC representation generates less runoff and more ET than the BCSD (and Maurer) VIC projections across much of the United States. In the Columbia

---

[1] Regridding done using CDO's remapcon operator. This is a first order conservative remapping.

15

River basin, however, there is a decrease in ET in LOCA-VIC relative to BCSD-VIC, with some corresponding increases in runoff. In the Canadian portion in particular, the large decrease in precipitation in LOCA relative to BCSD leads to a lower ET, runoff, and SWE in VIC output, this is attributed to the substantial differences in this region between Maurer and Livneh observation datasets.
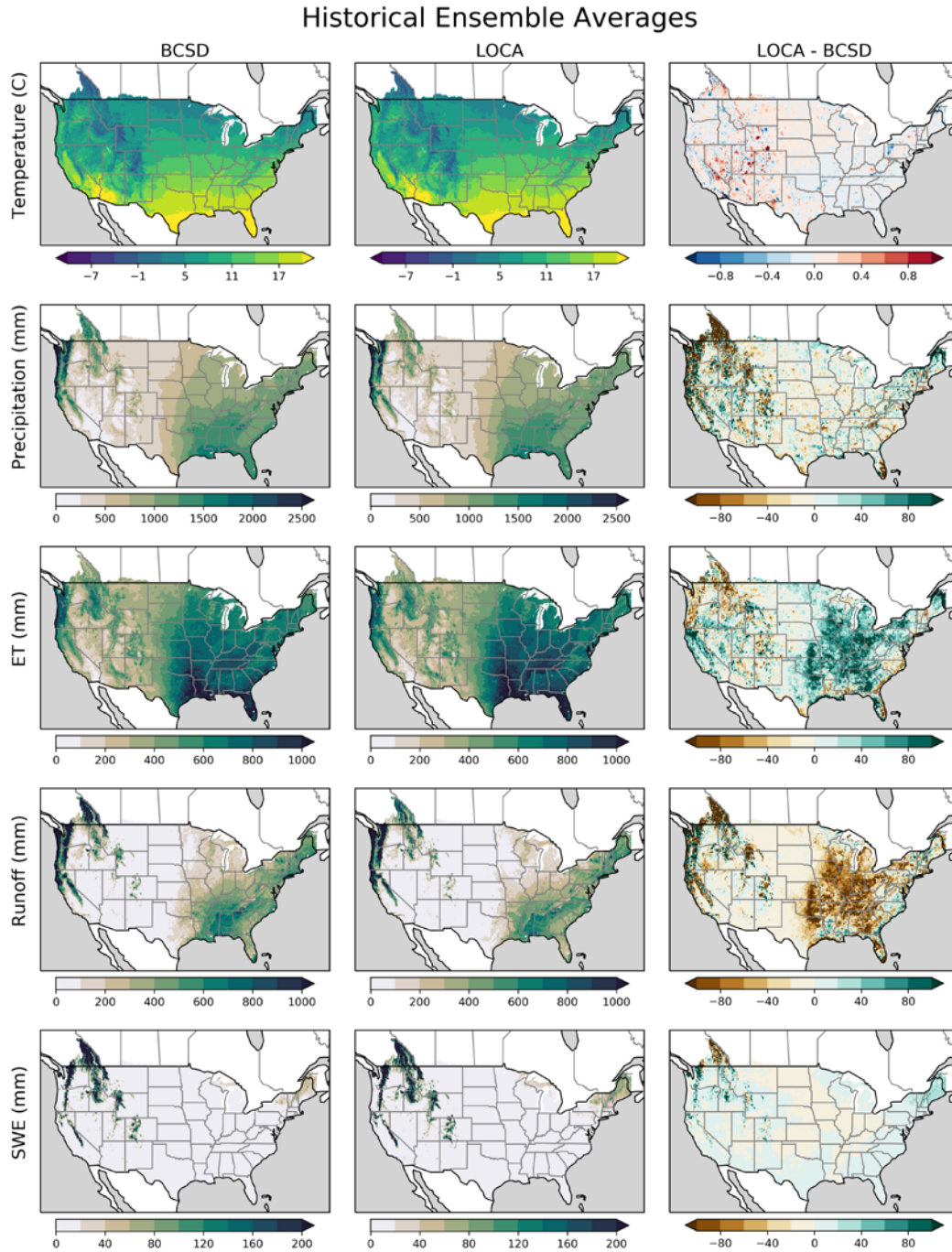

Figure 1 Ensemble average (23 GCMs) and differences between BCSD and LOCA for five hydroclimate variables. BCSD (left column) and LOCA (middle) and their differences (right). Historical period is 1970-1999.

Figures 2 and 3 provide an overview of average change signals for BCSD and LOCA for two emissions scenarios. Generally, temperatures increase everywhere for both methods, with changes between downscaling methods being mixed throughout the country (the change surface is less smooth in LOCA due to the analog selection process in LOCA and BCSD's reliance on the interpolated GCM temperature signal). Across both methods, precipitation tended to increase in the higher latitudes and decline in the Southwest. BCSD typically has greater increases and less pronounced decreases than LOCA, particularly in the western mountains, which is a result of the quantile mapping bias correction used in BCSD (as discussed in section 3.1.6). ET follows a similar pattern in both datasets, with less of a consistent difference between the two, though ET also typically has more of an increase and less of a decrease in BCSD compared to LOCA. Runoff declines in the mountains and in places of decreased precipitation generally in both BCSD and LOCA. However, in some mountain ranges in Wyoming and Canada, substantial increases in precipitation in BCSD result in increases in projected runoff. SWE declines everywhere, most prominently in the mountains, with LOCA having greater SWE declines in the East and BCSD having greater declines in the western portion of the domain, especially the Canadian portion of the Columbia.

Overall, change signals and differences between downscaling are amplified with greater emissions – i.e., change signals and differences for RCP 8.5 (Figure 3) are greater than for RCP 4.5 (Figure 2).
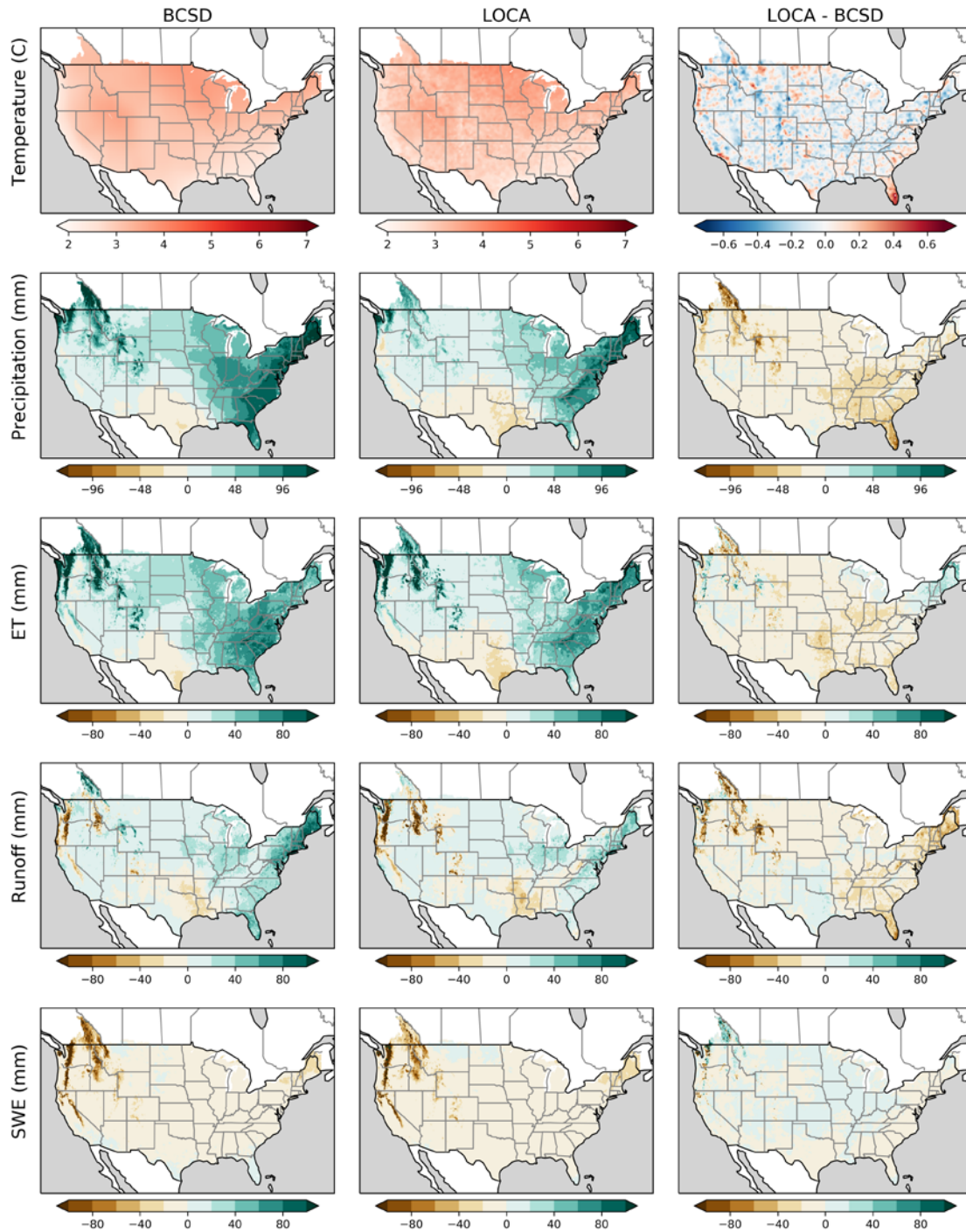
RCP 4.5 Change Signal

Figure 2 Ensemble average (23 GCMs) change signals for RCP 4.5. Changes between the historical period (1970-1999) and future period (2070-2099) for an ensemble of datasets generated with 23 GCMs using RCP4.5 emission trajectories.Plots in the left column are differences from datasets downscaled

using BCSD, the middle column are differences downscaled using LOCA, and the right column are differences (LOCA – BCSD) of the change signal between the two.
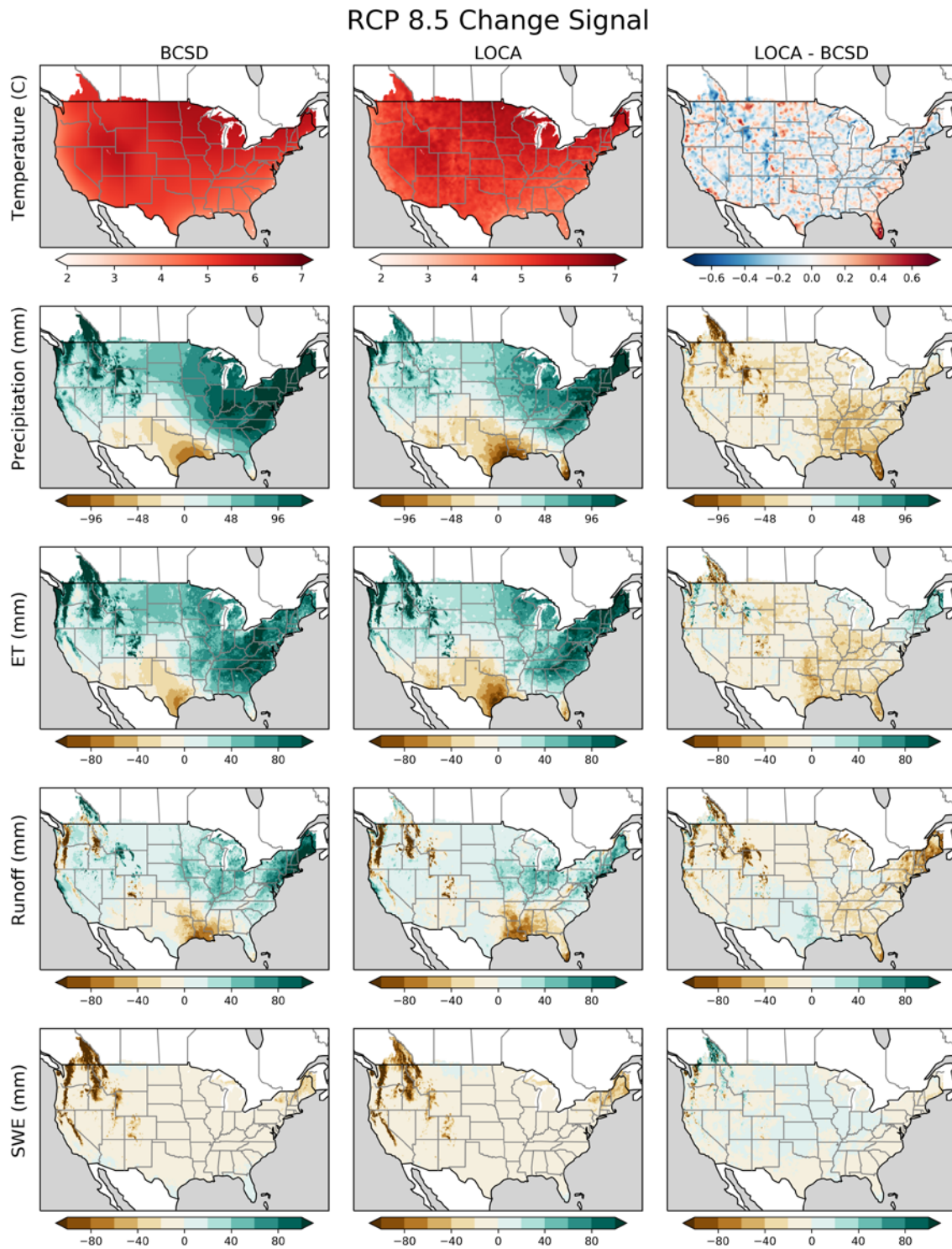

Figure 3 Ensemble average (23 GCMs) change signals for RCP 8.5. Similar to Figure 2, but for RCP 8.5 instead of 4.5.

3.1.2. Historical meteorological values (temperature, precipitation)

For 30-year mean annual temperatures and precipitation, both methods demonstrate similar patterns across the CONUS in comparisons to historical simulations from GCMs in both the ensemble average and the observational datasets and for three individual models (Figures 4 and 5) and all 23 models (not shown). On a continental scale, for long-term averages, both methods demonstrate fidelity. The observed dataset "OBS" on the top row of Figure 4 and 5, are indistinguishable from the ensemble averages in Figure 1, and downscaled output from individual GCMs show little difference relative to each other (see Figures A1, A2).

Difference plots (Figures 4, 5) between the observations and individual models make the differences that do exist more apparent (whereas, Figures A1a and A2a magnitude plots look identical). The differences between GCM historical simulations and the observational datasets of Maurer or Livneh (rows 2-4) demonstrate similar patterns across downscaling methods (comparing plots between the left and middle column within a row in Figures 4 and 5) than across GCMs (comparing plots within a column). We expect these plots to have differences because of natural variability – but difference plots allow us to see the nature of these differences better than total magnitude (shown in Appendix A) and how different they are in LOCA and BCSD. For temperature in particular, the colorscale is stretched such that very small differences, much less than 1°C, are visible (Figure 4). The differences between downscaling methods (right column) are consistent across individual GCMs and also appear in the differences between observational datasets (top row, right column). The high degree of similarity between these LOCA-BCSD plots, and between these plots and the top row right column (differences between observation datasets) illustrate the influence of the observational datasets in downscaling method comparisons. The one notable departure from this pattern is the Canadian portion of the Columbia River basin. In this basin, LOCA is consistently drier than the Livneh observations across all GCMs, and BCSD is wetter than the Maurer observations across all GCMs (Figure 5).

The consistent differences between the LOCA and BCSD datasets appears as the fine-scale polka dot like features surrounding individual weather stations that were presumably treated differently in the Livneh and Maurer datasets (or absent in one). This polka dot correspondence is present in both the temperature dataset (Figure 4) and the precipitation dataset (Figure 5). There are particularly exaggerated differences in southern Florida and throughout the mountain west in which the dots blend together. Note the polka dots are in the exact same place and relative magnitude in the top row comparing Maurer and Livneh datasets as they are in the lower rows comparing BCSD and LOCA datasets. This is an important point because many of the following differences between datasets likely stem from the differences between the Maurer and Livneh datasets themselves rather than the downscaling method. Users should understand how their system would be represented when using either the Maurer or Livneh dataset as "observations" before trying to evaluate the expected effect of LOCA or BCSD dataset in their analysis. A climate change analysis should never be performed comparing a projected future climate from one product (e.g., LOCA) from the historical representation of another (e.g., Maurer). Note at times there are squiggly lines across BCSD plots related to the HUC2 boundaries; BCSD was implemented independently for each HUC2 and as a result there can be artifacts on the boundaries in a continental domain analysis (discussed again in section 3.4).

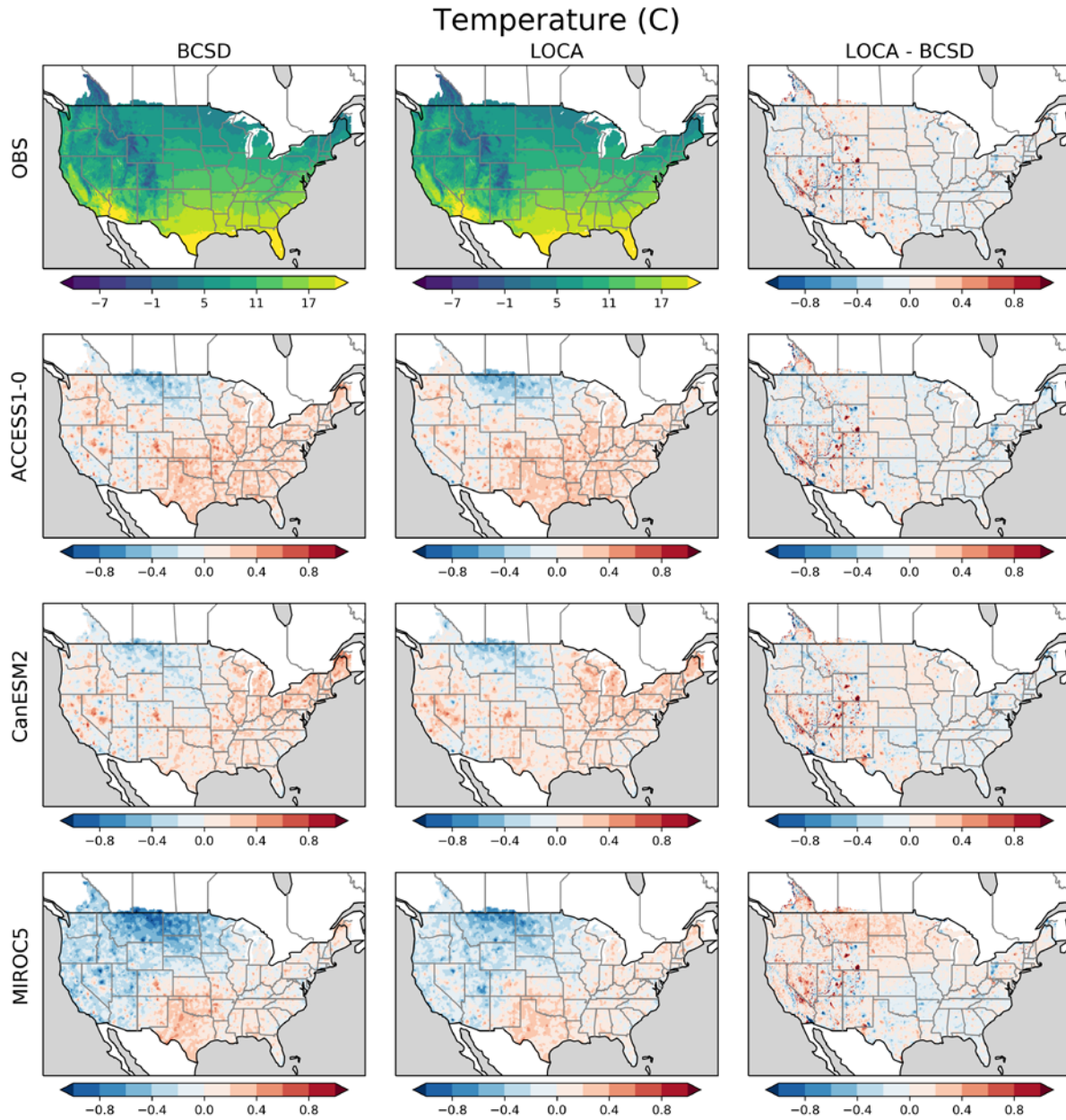For information on interannual variability, see Appendix B.

Figure 4 Historical mean annual temperature differences. Observations (top row) for both BCSD (left column, Maurer) and LOCA (middle column, Livneh). The three individual model examples (rows 2-4) are differences between the modeled historical period and observations specific to the dataset used by the downscaling method. The third column is the total difference (LOCA – BCSD) for observations (row 1) and downscaled GCMs (row 2-4). All values are averages from 1970-1999.
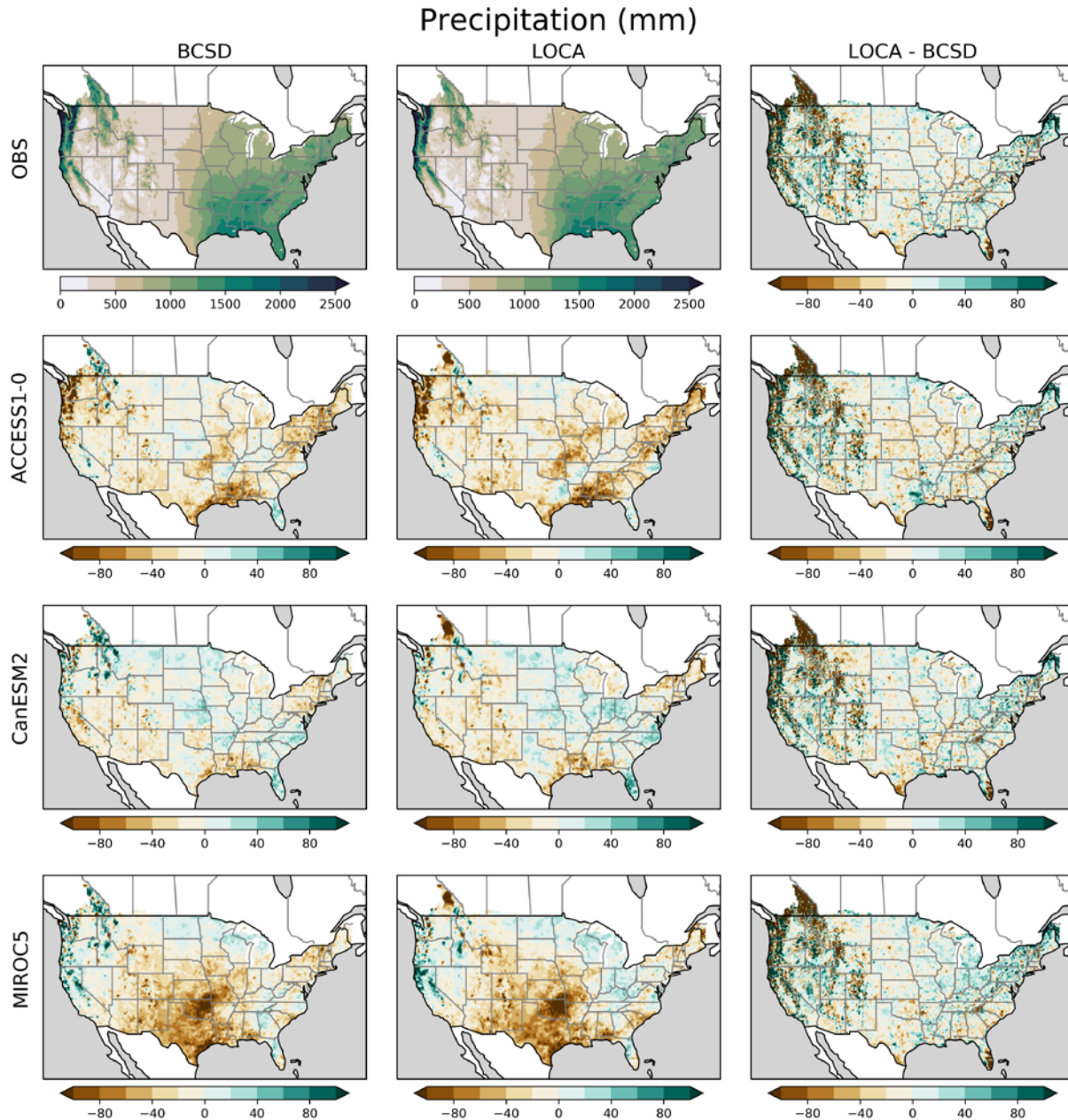
Figure 5 Historical mean annual precipitation differences. Identical format to Figure 4, except for precipitation instead of temperature (for percent differences, see Appendix A).

3.1.3. Historical hydrological values (ET, runoff, SWE)

We explore comparisons using the hydrologic variables ET, runoff, and SWE computed with VIC driven by BCSD and by LOCA for the CONUS domain. The VIC hydrologic model versions used differed (BCSD used 4.1.2, LOCA used 4.2.c) but differences between these two versions have been evaluated and are small, especially relative to other factors in this comparative evaluation. The more substantial differences in hydrology are driven by observational dataset differences and their respective grid sizes.

As with temperature and precipitation, 30-year annual mean ET, runoff, and SWE, hydrology as evaluated over the historical period compares closely to hydrology forced with the historical

datasets (Maurer and Livneh respectively). On a continental scale, for long-term 30-year averages, both methods show little difference relative to each other (i.e. all figures appear similar, see Appendix A). However, on a finer scale, important differences exist (Figures 6, 7, 8). In particular, the higher-resolution of LOCA (6 vs. 12 km) results in mountain-top grid-cells with more precipitation and colder temperatures and thus more average SWE (Figure 8), which can sustain snowpack and runoff later into the summer.

Difference plots (Figures 6, 7, 8) between VIC simulations run with observations and those run with downscaled historical GCM output are used to illustrate the differences that are present. As mentioned before, we would not expect these two representations of the historical period to be identical because of natural variability. As with temperature and precipitation, hydrologic simulation differences relative to their respective observation dataset are smaller between downscaling methods (within the same row) than differences between historical GCM simulations (within the same column). Larger differences exist between VIC simulations run with the Maurer dataset and simulations run with the Livneh dataset (top right). Livneh-based simulations compared to Maurer show substantially more ET in the non-coastal eastern CONUS and less ET in the mountains of the Pacific and interior Northwest. Differences between LOCA-BCSD for simulations run with individual downscaled GCMs are largely similar to differences between the observational datasets (right column). Again, this illustrates the influence of the observational dataset used to train the statistical downscaling.

One of the factors that differs between the Livneh and Maurer observational datasets is a substantial increase in wet day fraction observed in the Livneh dataset. This increase stems in part from the increased number of grid cells which require interpolation between station observations in Livneh's 1/16th degree dataset (see section 3.3.4 for a more in-depth look at wet day fraction). More wet days have competing effects: first they result in more light precipitation events for which precipitation is more likely to remain in the vegetation canopy or soil surface and evaporate before reaching the sub-surface. However, wet days also result in an effective decrease in solar radiation to the surface because they imply the presence of clouds. In VIC simulations in particular, the MTCLIM [*Thornton and Running,* 1999] algorithm used to generate forcing variables for VIC estimates the presence of clouds based on the occurrence of precipitation, and thus decreases shortwave radiation more in datasets with higher wet day fraction. Especially in dry areas and months, the BCSD method occasionally needs to handle the case where the selected analog month experienced no precipitation, but the model month being downscaled did. In this event BCSD distributes the model precipitation across all days in the month (though not uniformly), yet this yields a month with all wet days, which would affect the MTCLIM solar radiation estimates and VIC's evapotranspiration estimates in such locations. This approach differs from the original implementations of BCSD, described in *Harding et al.* [2012]. To avoid such wet-day impacts, BCSD was designed to find appropriate resampling patterns in an iterative fashion when the sample month lacked adequate wet-days to support the desired downscaled precipitation total.

As with ET, the runoff simulated by VIC differs more when comparing the two observation datasets then either of the downscaling methods differ from the observation-based simulations (Figure 7). The differences across GCMs (looking down a column) are greater than the differences between the downscaling method in comparison to their respective observation-based simulation. Also, the differences between LOCA and BCSD derived runoff (right column) are very similar to the differences between the two observation-based simulations regardless of the GCM used.

The large differences in runoff between simulations likely results from the change in precipitation patterns in the Livneh and Maurer datasets. The Livneh dataset has lower intensity precipitation compared to Maurer, especially in high elevations in the West (see section 3.3.1 on Annual Maximum precipitation for more details). This could be due to interpolation requirements between the datasets operating differently at their different grid spacings. The decrease in intensity results in a decrease in runoff over much of the CONUS domain. There are also very large differences in the Canadian portion of the Columbia River basin (also shown in figures in *Livneh et al.* [2014]), most likely due to a change in the stations and the reference climate normal used [*Livneh et al.* 2015].

Last, differences between LOCA and BCSD VIC modeled SWE are presented in Figure 8. As with other variables, there is a consistency between the observational based differences and the downscaled climate differences. Across all datasets, LOCA-VIC simulates a greater mean SWE in the mountain ranges in the western CONUS and in New England compared to BCSD-VIC. In New England and the western CONUS, this difference matches the differences between Livneh-VIC and Maurer-VIC and likely stems from subtle differences in seasonal precipitation and temperature, for example, there is slightly more precipitation in New England in Livneh than in Maurer (Figure 5). However, in the Canadian portion of the Columbia, LOCA-VIC models substantially less SWE than BCSD-VIC. This comes from three places: Livneh-VIC has less SWE than Maurer-VIC in this region (brown in top right figure). Second, LOCA-VIC has a little less SWE than Livneh-VIC (brown in middle column difference plots); finally, BCSD-VIC has a little more SWE than Maurer-VIC (green in left column difference plots).

Overall, the majority of differences are apparent in observational datasets – as is demonstrated by the top row of Figures 4, 5, 6, 7, and 8 being very similar to the ensemble averages and differences in Figure 1.
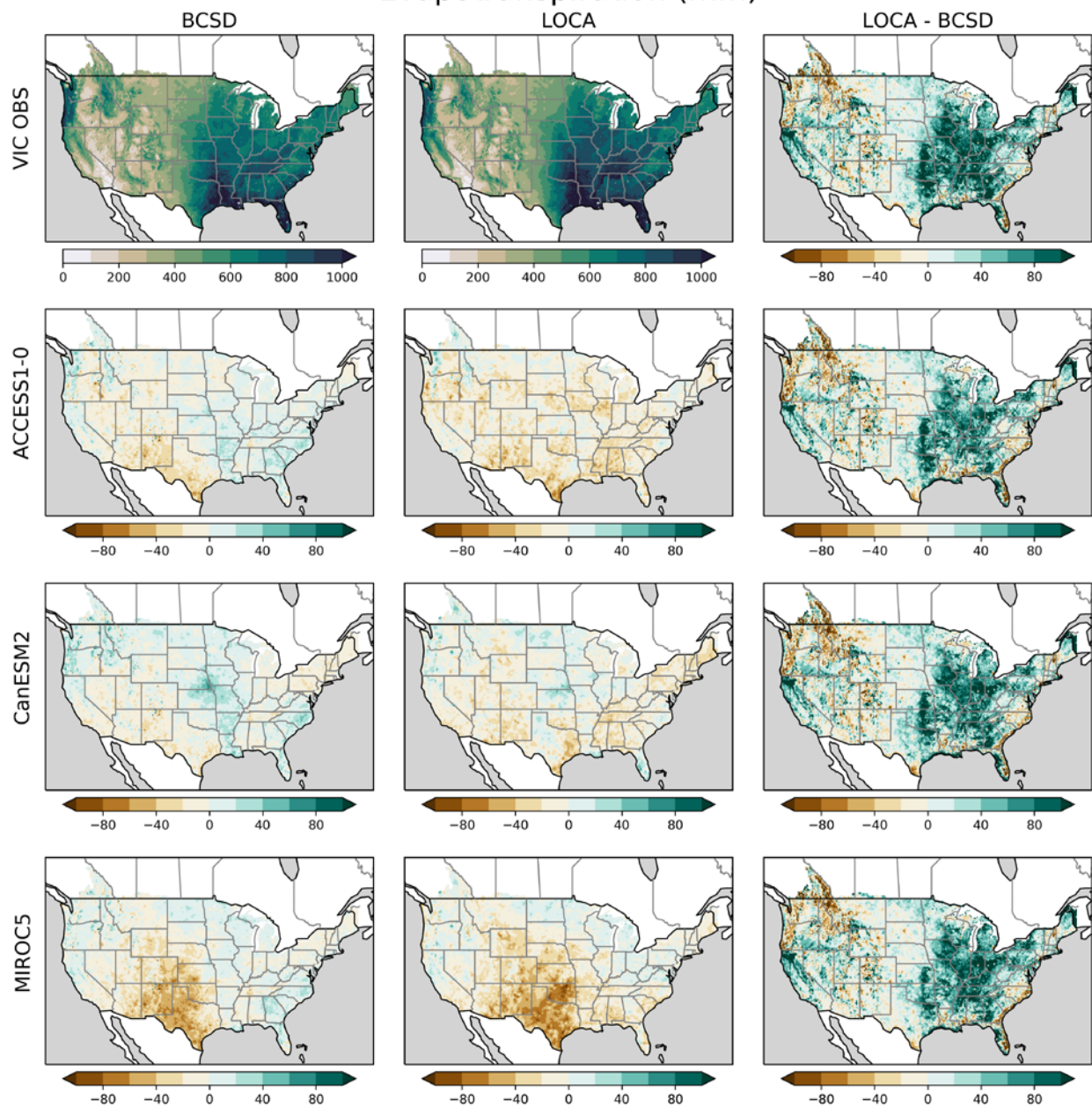
Evapotranspiration (mm)

Figure 6 Historical mean annual evapotranspiration differences. Identical format to Figure 5, except for ET instead of precipitation.
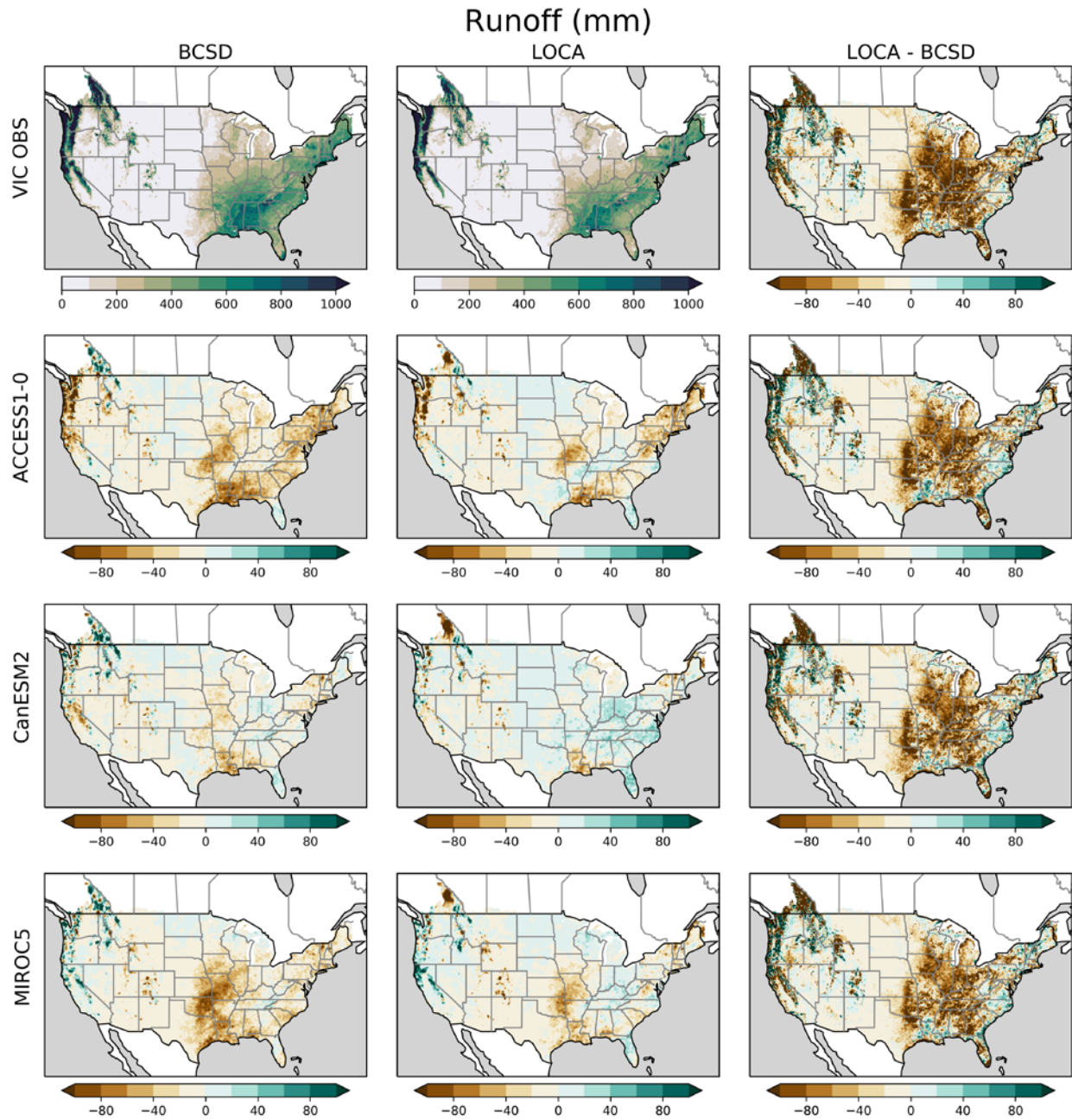
Figure 7 Historical mean annual runoff differences. Identical format to Figure 5, except for runoff instead of precipitation.
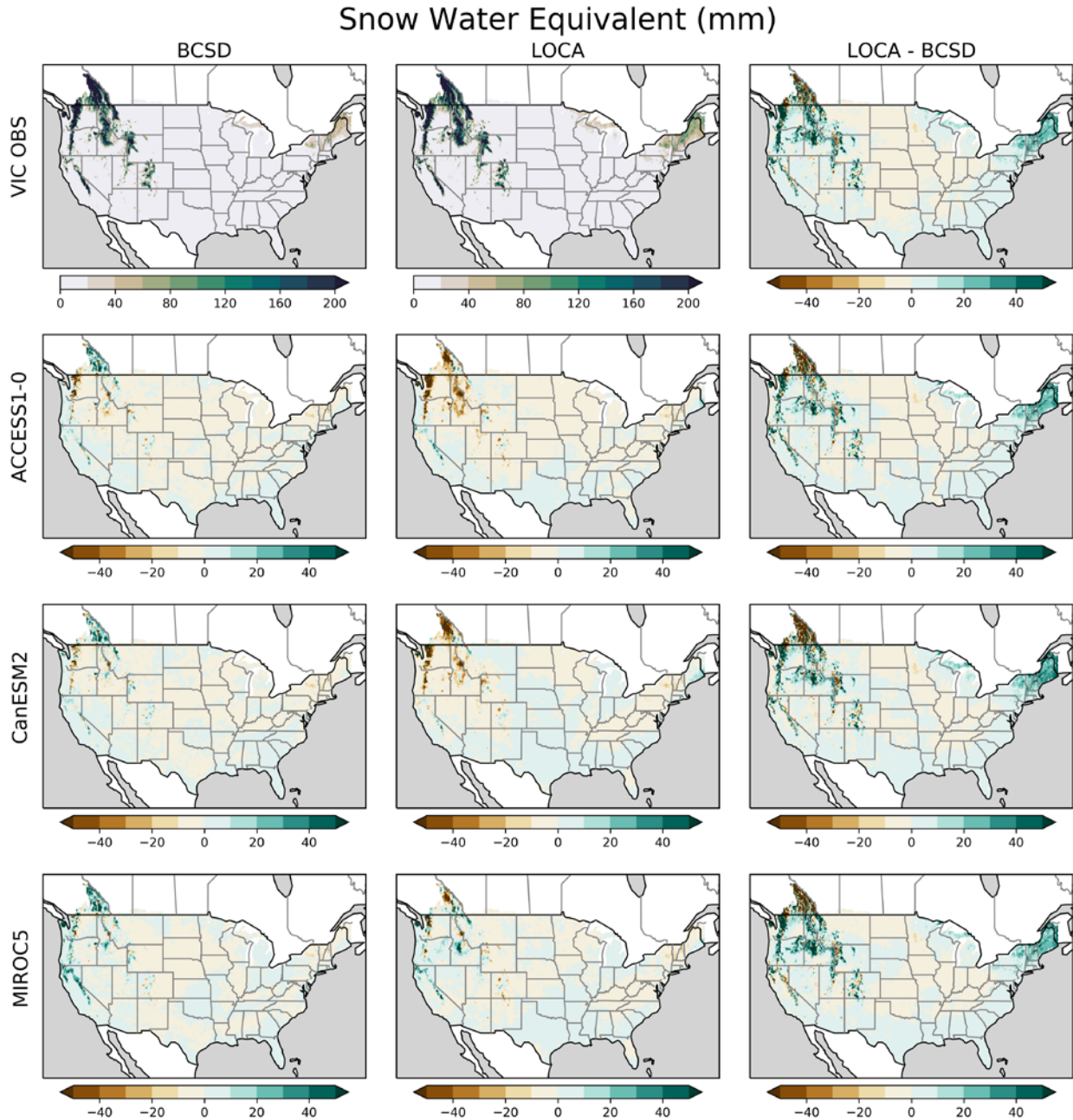
Figure 8 Historical mean annual SWE differences. Identical format to Figure 5, except for SWE instead of precipitation.

3.1.4. Future climate conditions in meteorological values

In addition to historical comparisons, we evaluated the future climate changes in each of the five variables (temperature and precipitation in this section; ET, runoff, and SWE in section 3.1.5). Figures below show changes for LOCA and BCSD between the historical (1970-1999) and future period (2070-2099) for RCP 4.5 and 8.5 (Figure 9 and 10). The magnitude of these changes (right column) are displayed next to differences between RCP 4.5 and 8.5 (bottom row) for context. The influence of downscaling method on the magnitude of changes in future projections is most apparent when averaged across ensembles (right column in Figures 9 and

10). We also show changes over a subset of individual GCM responses (see figures in Appendix C). As mentioned earlier, we show these to reinforce the importance of recognizing how individual model responses differ from the ensemble average.

Future temperature projections show consistent warming across both downscaling methods, with RCP 8.5 (high emissions) scenario being about 2-3°C warmer than RCP 4.5 (medium emissions) scenario (Figure 9). Differences between downscaling methods are much less than differences between RCPs. BCSD has smoother changes than LOCA, resulting in locations being hotter and others being drier throughout the country, though notably, these differences are relatively small. The spatial heterogeneity in LOCA is a product of the observation dataset, the analog selection process in LOCA, and BCSD's use of the interpolated GCM changes for temperature. Note also that the broad spatial patterns for a given GCM are consistent between methods. For example, both BCSD and LOCA show more warming over the regions that are represented in the GCM with higher elevation terrain relative to the surrounding region in MIROC5 (and ACCESS1-0 to a lesser extent), see for example the warmer region around 40N, -110E in Figure C1. This is most likely due to an elevation dependent warming in the GCM that follows the GCM terrain, caused by the snow albedo feedback and other processes.

Future precipitation projections across downscaling methods (Figure 10, C2) show similar patterns in change signals with the higher latitudes becoming wetter and lower latitudes drier, though LOCA is consistently drier than BCSD (brown areas in the right column of Figure 10). This is shown with individual GCMs as well (Figure C2), though the pattern is less clear because increases are interspersed. Both downscaling methods show similar differences between RCP 4.5 and 8.5 (bottom row, Figure 10). These differences are similar in magnitude to differences between downscaling methods (right column, Figure 10).
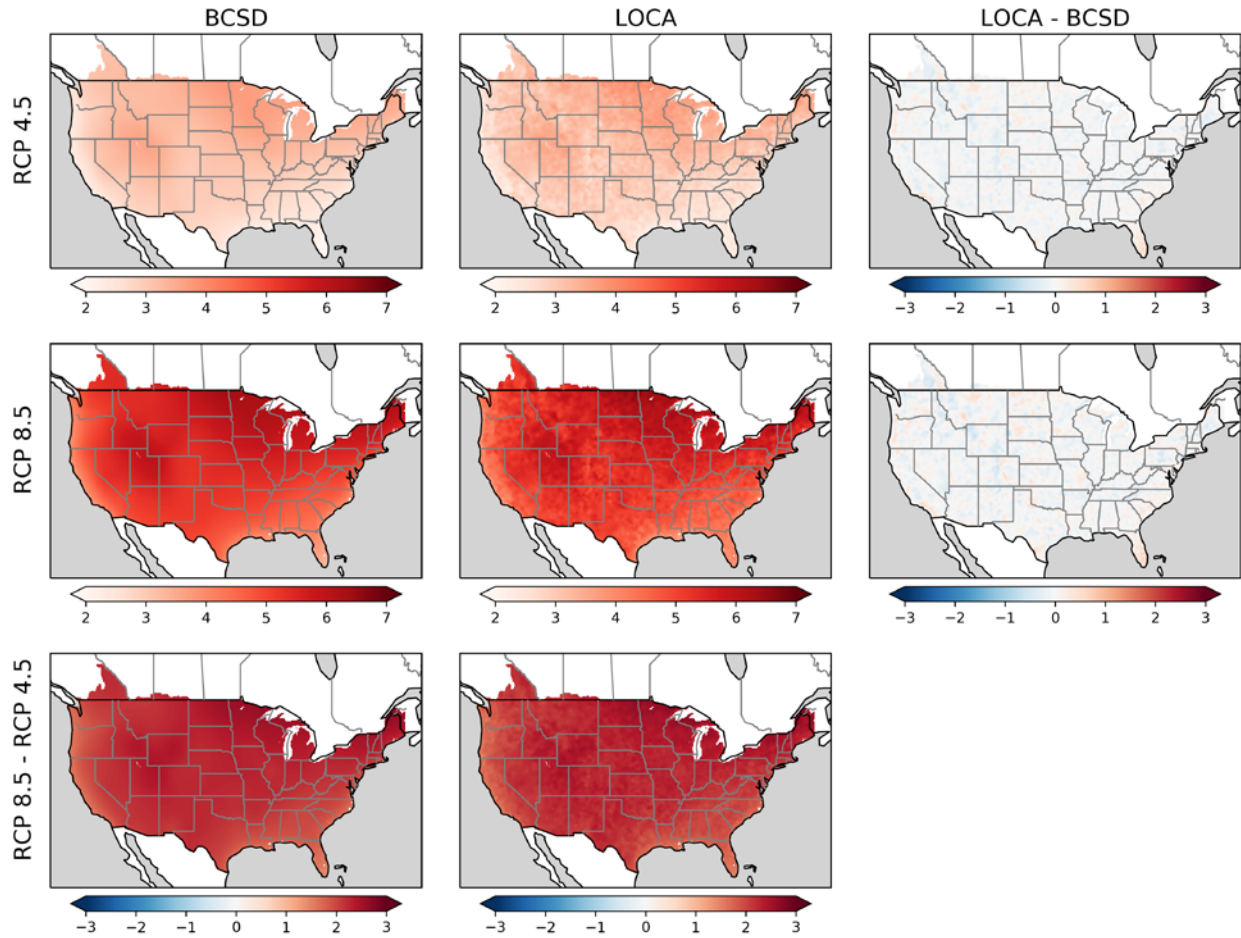
Figure 9 Changes in annual temperature for 23-model ensembles. Differences between future changes (2070-2099) and historical (1970-1999) for the ensemble average for RCP 4.5 (top row) and RCP 8.5 (middle row) for both BCSD (left column) and LOCA (middle column). The third column is the total difference (LOCA – BCSD) between downscaling methods. The bottom row is the total difference between RCPs.
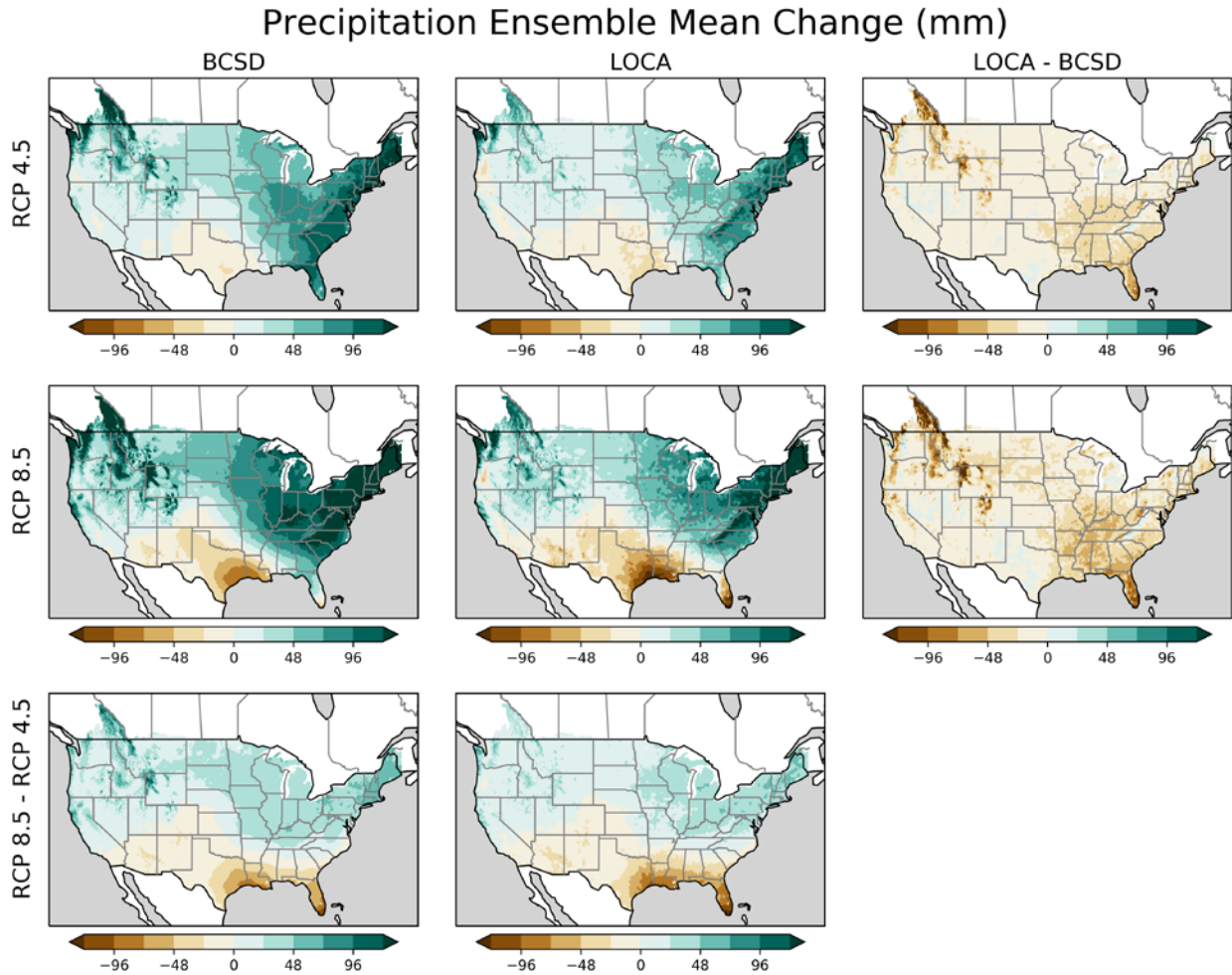
Figure 10 Changes in annual precipitation for 23-model ensembles. Format similar to Figure 9, but for precipitation instead of temperature.

3.1.5. Future climate conditions in hydrologic values

For hydrologic variables, the magnitude of the climate change signal is influenced by downscaling methods and their associated hydrologic model configuration. The extent of these changes, however, depends on the particular location and variable of interest. As with changes in meteorology, there are more similarities between LOCA-VIC and BCSD-VIC for the same GCM than from LOCA-VIC (or BCSD-VIC) for different GCMs.

For future projections of ET (Figure 11, C3), on average, both BCSD-VIC (left column) and LOCA-VIC (middle column) show similar patterns: ET increases in most locations except for in the lower latitudes where precipitation is also decreasing. These changes are larger with higher emissions (i.e., RCP 8.5 changes are greater than RCP 4.5), but the spatial patterns are the same. These changes generally mimic the changes present in precipitation (Figure 10), with some modulation by local land surface characteristics, e.g., the southern Mississippi River. LOCA-VIC projects less of an increase and more of a decrease than BCSD-VIC as indicated by the dominance of brown colors in the LOCA-BCSD plots (Figure 11, right column), though exceptions appear in the Northeast and some locations in the western mountains. Individual models (Appendix C) show considerably more variability; however, the ET response to climate change differs more depending on GCM (differences between rows) than it does between

downscaling methods (differences between left and middle column for the same row), and the change signal mirrors the change signal in the precipitation signal in each individual GCM (Appendix C). One notable exception to the tight coupling between precipitation change and ET change is in the upper Columbia River basin. In this region, ET is more energy limited than in other regions, as such the change in precipitation has less of a direct effect on ET than the change in temperature does. For example, with ACCESS1-0, LOCA projects a decrease in precipitation, while LOCA-VIC projects an increase in ET. While this is the most notable example, similar patterns can be seen across higher latitude mountains. This is also evident in runoff as will be discussed next.
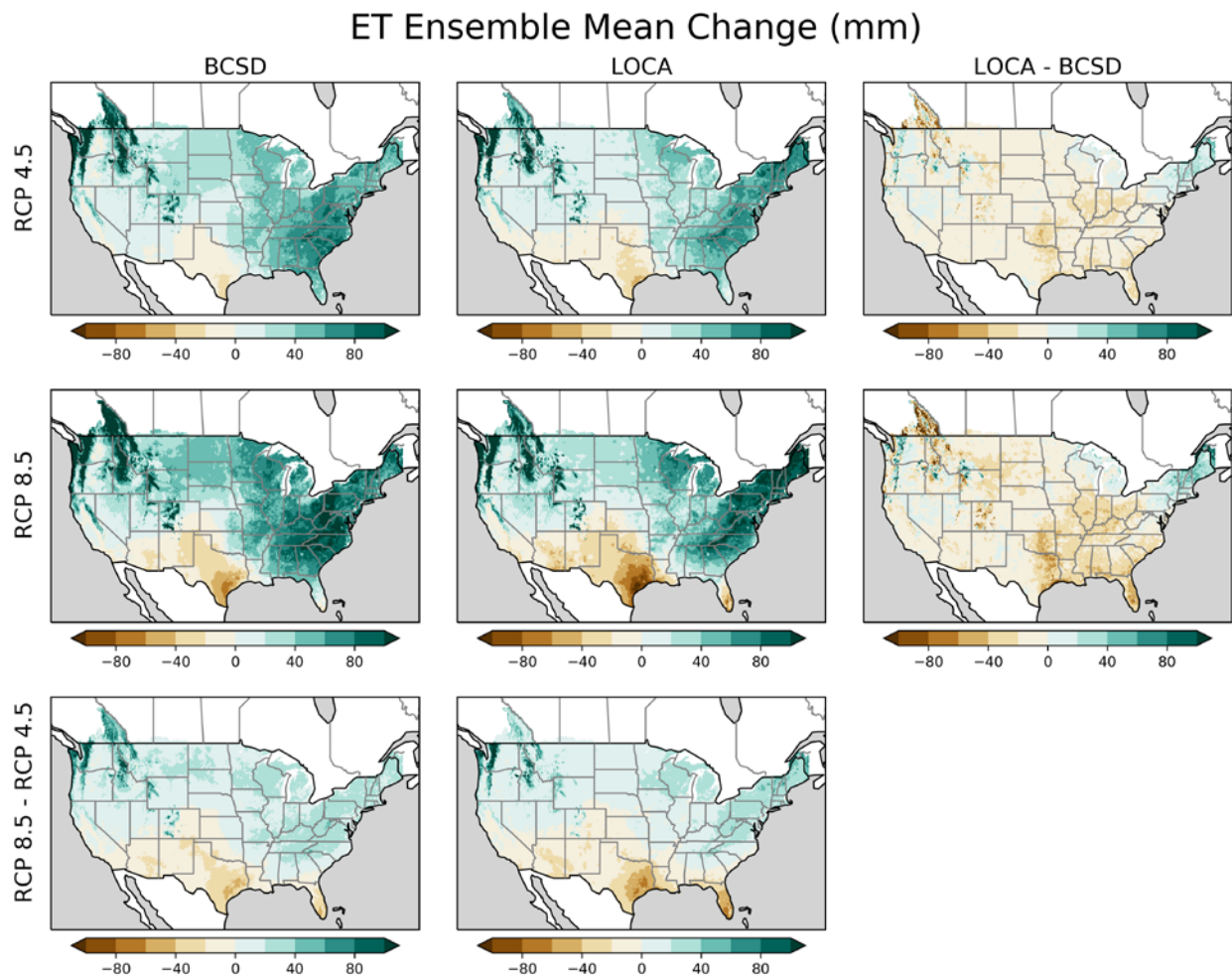

Figure 11 Changes in annual evapotranspiration for 23-model ensembles. Differences between the historical period (1970-1999) and future period (2070-2099) for RCP 4.5 (row 1) and 8.5 (row 2) for BCSD (left) and LOCA (middle) and the differences between LOCA and BCSD (right). Row 3 is the difference between RCP 8.5 and 4.5.

For runoff (Figure 12 and c4), again, both downscaling methods have similar patterns (left and middle columns) with notable exceptions in the western mountains. LOCA-VIC has less increase in runoff than BCSD-VIC in the East. Moreover, considerable variability exists in the western mountains, with a more consistent decrease in future runoff in LOCA-VIC, particularly in the upper Missouri and upper Snake river basins (Montana, Idaho). The relatively sharp border between the CONUS and Canada in the Columbia River basin is concerning, with similar effects in both LOCA and BCSD, though this is less pronounced in BCSD. The cause of this sharp

border likely stems from changes in the statistics of precipitation products in this region due in part to differences in precipitation measurement networks, and possibly due to changes in the VIC parameters across the border.

Individual GCMs (Figure C4) generally show less increase for LOCA-VIC and more decrease, relative to BCSD-VIC (more browns in right column), but here there is some variation across GCMs, particularly in the southern half of the United States. The Columbia River basin (inland of the first coast range), much of the interior mountain ranges in the western CONUS, and New England all have consistently less runoff in LOCA-VIC than BCSD-VIC. This is consistent with the patterns noted in ET, in which energy limited environments have increases in ET regardless of changes in precipitation. In addition, the smaller increases in precipitation projected by LOCA in these regions means that the increase in ET overwhelms the increase in precipitation, resulting in a net decrease in runoff. In BCSD, there is a much larger and more consistent increase in precipitation in these regions that arises from the quantile mapping bias correction (see section 3.1.6.) as a result, the increase in ET is not sufficient to counter the increase in precipitation, and there are some increases in runoff (also noted in *Lukas et al.* [2014] and *Reclamation* [2011]). Because the changes in temperature in both BCSD and LOCA are very similar, the increase in ET in these environments is similar in both LOCA and BCSD, and as a result, the smaller increase in precipitation in LOCA translates into a decrease or smaller increase in runoff.
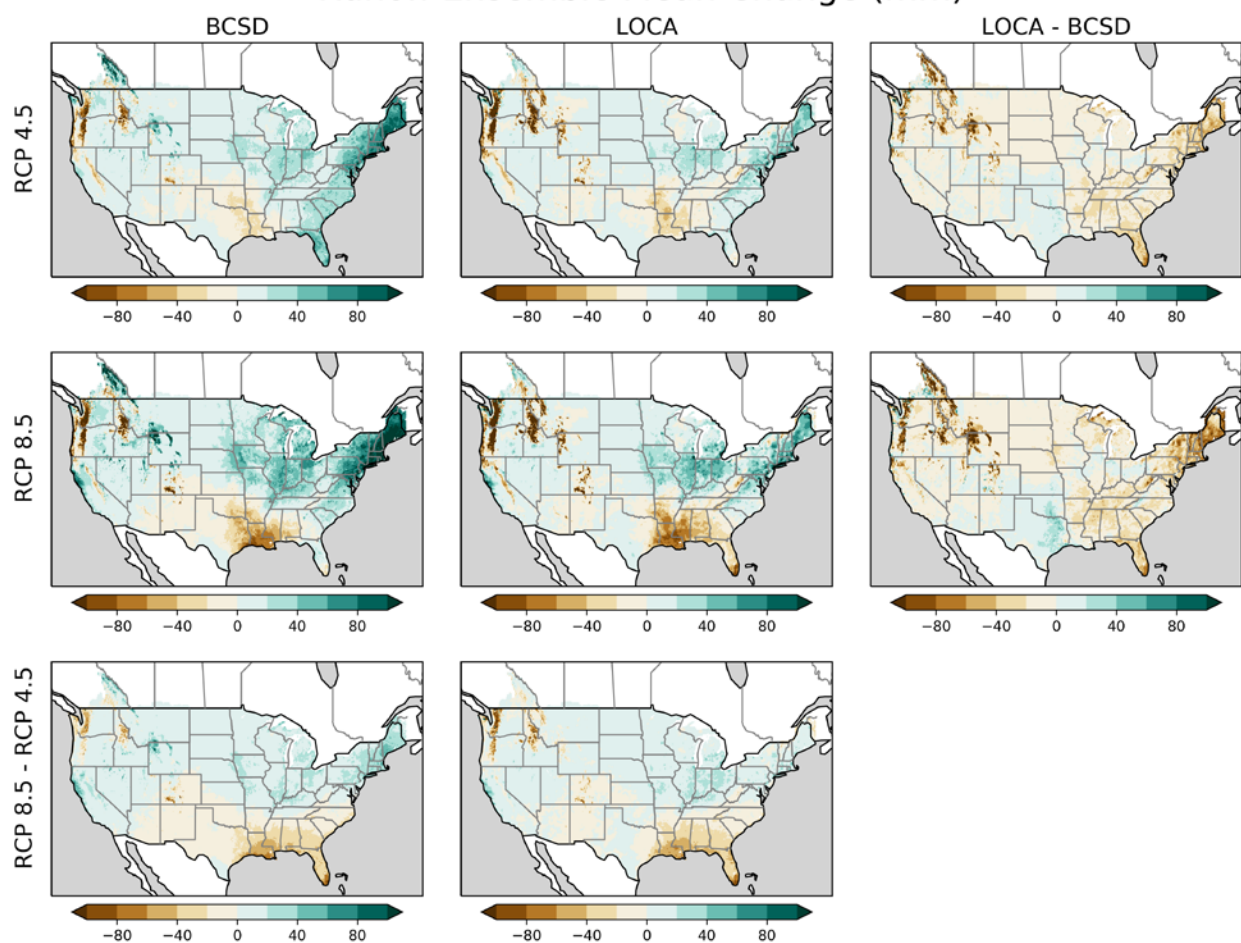
Runoff Ensemble Mean Change (mm)

Figure 12 Changes in annual runoff for 23-model ensembles. Similar format to figure above but for runoff (surface runoff + baseflow) instead of ET.

Future projected changes in SWE (Figures 13 and C5), show declines across the CONUS, with sharper declines in RCP 8.5 compared to RCP 4.5. SWE changes are similar in LOCA-VIC and BCSD-VIC, with exceptions again in the western mountains. Some mountain locations have greater SWE with LOCA-VIC and others with BCSD-VIC, and these locations are not always consistent across GCMs (right column), with the exception that in the Canadian portion of the Columbia River basin and the northern interior mountain ranges of the CONUS, LOCA-VIC projects less of a decrease in SWE than BCSD-VIC (Figure 13 and C5), particularly in RCP 8.5. This may be due to higher spatial resolution in Livneh and LOCA. This higher spatial resolution permits the Livneh dataset, and thus LOCA, to resolve the colder mountain tops better, and results in more areas for which the warmer temperatures alone are insufficient to cause as much of a decrease in SWE. The VIC model represents sub-grid variability through the use of multiple snow elevation bands within a grid cell, so the link is not as direct as it would be in other hydrologic models; however, these bands feed into a single common soil column, so it is not a complete hydrological representation of sub-grid elevation.
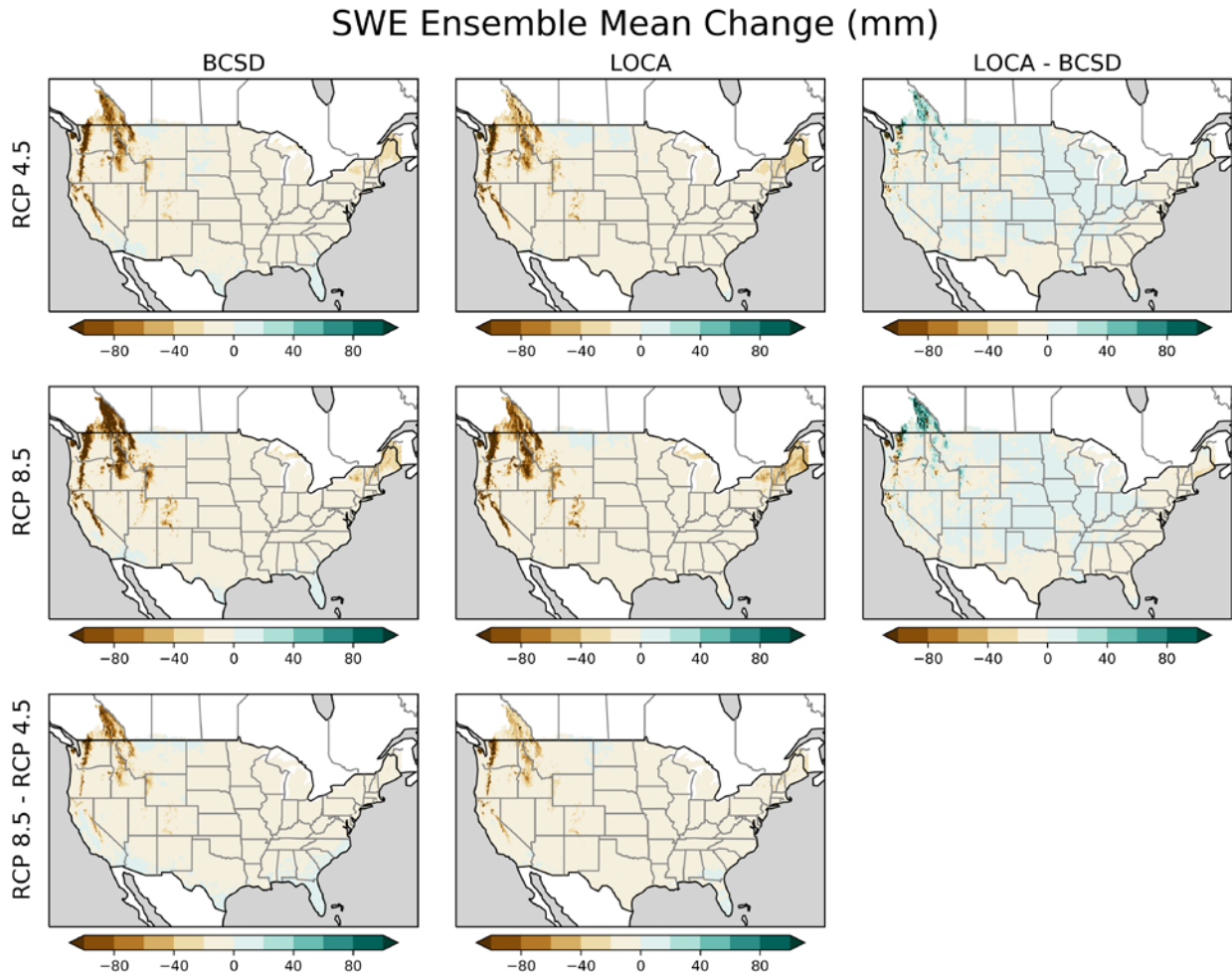
SWE Ensemble Mean Change (mm)

Figure 13 Changes in annual SWE for 23-model ensembles. Similar format to figure above but for SWE.

3.1.6. Explanation of bias correction differences

BCSD and LOCA methods treat bias correction differently, resulting in different precipitation projections seen between the two datasets in Figs. 2 and 3. GCM-projected precipitation changes by the end of the century show increases across much of the northern part of the CONUS and decreases in Texas and the Southwest, as shown in the left column of Figure 14 for RCP 8.5. The middle column shows the difference between the BCSD downscaled result and the original CMIP5 result, while the right column shows the analogous difference for the LOCA downscaled result. The LOCA process attempts to preserve the original CMIP5-predicted precipitation change, while the BCSD bias correction method allows it to be altered. As described in *Maurer and Pierce* [2014], the quantile mapping bias correction used in BCSD changes the long-term (multi-decadal) climate change signal seen in the CMIP5 dataset depending on how the monthly GCM variability compares to observations, since the short timescale variability is stronger and quantile mapping draws no distinction between variability on short and long timescales. The result is that BCSD shows an overall shift towards wetter projections in most of the western half of the CONUS and parts of the Eastern CONUS Changes in runoff are about twice the magnitude in precipitation in most of the Southwestern CONUS (i.e., the precipitation elasticity is about 2 [*Sankarasubaramanian et al.,* 2001]), so a

34

~2-8% shift towards wetter projections as seen in BCSD could shift the original CMIP5 runoff projections by ~4-16%.

There is not a consensus on the degree to which bias correction and downscaling approaches need to capture the raw GCM change signal.   Some would argue that statistical downscaling should preserve the GCM signal, because it is based on physics, while others would argue that the GCM change signal is based on very low-resolution representation of those physics and there should be no expectation that changes in, e.g. orographic precipitation, or local air temperature due to land-atmosphere interactions should remain the same as in the original GCM. A review of some relevant bias correction issues is given in *Maraun* [2016].
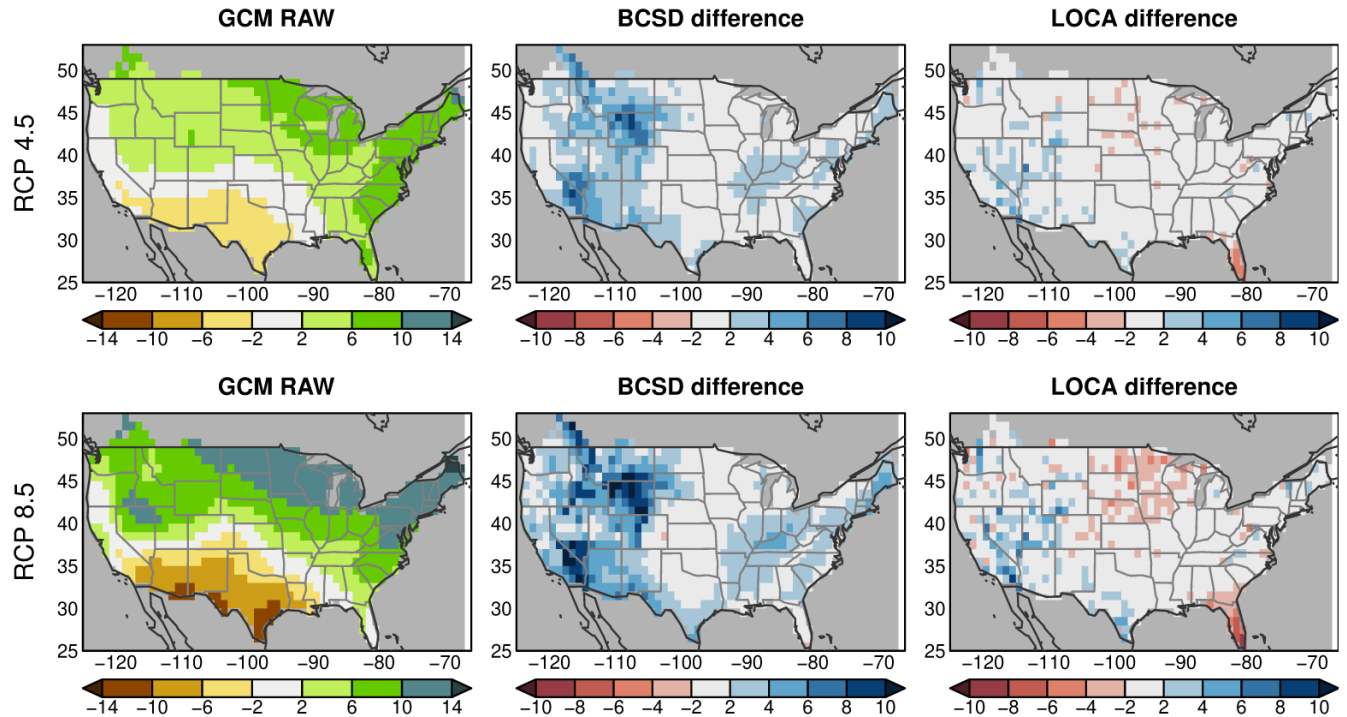

Figure 14 Changes in annual precipitation for 23-model ensembles. Model-projected changes (%) in yearly precipitation, 2070-2099 with respect to 1970-1999, in RCP 4.5 and 8.5 emissions scenarios. Left column: multi-model ensemble averaged results from the original 23 CMIP5 GCMs (raw). Middle column: difference (percentage points) between the LOCA result and the CMIP5 GCMs. Right column: difference (percentage points) between the BCSD result and the CMIP5 result. All fields have been bilinearly interpolated to a common 1°x 1° grid for computing the difference.

3.2 Seasonal changes

Seasonal changes in when water arrives and is in rivers and streams has considerable impact on society and ecosystems. To evaluate seasonal changes, we compare runoff averages within each HUC2 (Figure 15) in each month of the year. These differences are driven by a number of factors, though largely driven by seasonal precipitation changes, which we explore in more detail in section 3.2.3.

3.2.1 Historical hydrographs

Figure 15 Location of HUC2 watersheds. We provide six examples from HUC2s below. All 18 HUC2s appear in Appendix D.

When comparing seasonal hydrographs between methods (Figure 16), in general, the downscaled-based VIC datasets are more similar to their respective observational-based VIC dataset than they are to each other. In the left column, dark solid lines represent runoff simulated by VIC using historical observations from Livneh (dark blue) and Maurer (dark orange), and lighter lines represent the ensemble of historical GCMs run with LOCA (blue) and BCSD (orange). Differences between the dark and lighter line of the same color illustrate how close the GCM downscaled VIC ensemble mean matches observation-based VIC simulations.

In some locations (e.g., HUC 14), the historical observation-based datasets are relatively similar, while in other locations (e.g., HUC 5, 9, 11) they are very different. Most notable differences occur in the winter and spring. In most regions and months (exceptions in HUC 1 and 13) Livneh and LOCA-VIC have less runoff than Maurer and BCSD-VIC. The only HUC in which the downscaled VIC simulations are more similar to each other than they are to their respective observational-based VIC simulation is HUC 15. This may be a result of the difficulty in simulating both the hydrology and the meteorology in the desert southwest. In this region, GCMs do not simulate the North American monsoon well, and features that typically lead to the development of a local snowpack, e.g. the Mogollon rim, are not resolved either.
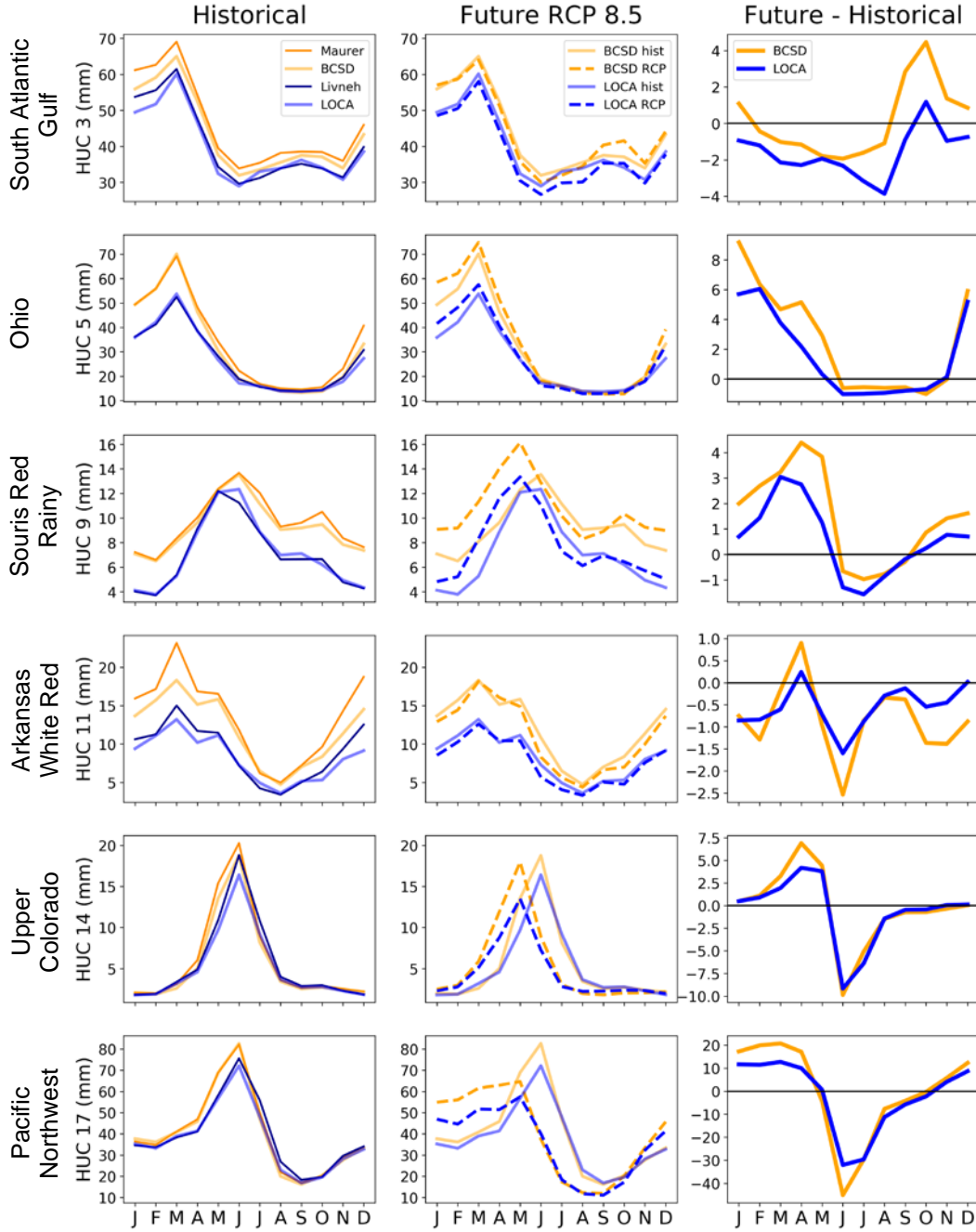
Runoff (mm)

Figure 16 Historical and future HUC2 basin runoff. Runoff (unrouted) aggregated by USGS HUC2 by month. Six examples above. For all 18, see Appendix D. The left column compares historical values from observations (dark line) and 23-ensemble GCM averages (lighter line) for 1970-1999. The middle column compares the same historical values (solid line) with future values of the same ensemble from RCP 8.5 from 2070-2099 (dotted line) for LOCA (blue) and BCSD (orange) downscaling. The right column shows the simulated climate change signal (i.e., differences between the historical and future ensemble means). Values are in mm, averaged across the watershed area.

3.2.2 Future changes in hydrographs

The changes in BCSD-VIC and LOCA-VIC runoff across HUC2s (Figure 16, middle and right columns and Appendix D) also shows seasonal patterns are similar between downscaling methods, particularly in snowmelt dominated regions. However, the magnitude of the change is highly variable. In some regions, (e.g., HUCs 10, 12, 14, 16, 17, 18) the climate change signal is relatively similar, but in some regions, the magnitude of the climate change signal differs substantially, and even changes sign between the LOCA-VIC and BCSD-VIC projections. For example, in HUC 3, the BCSD-VIC ensemble average projects an increase or very little change in discharge throughout the year, whereas the ensemble average projection in LOCA-VIC is a decrease throughout the year. In some cases, the differences between projections may be due to the higher resolution LOCA-VIC simulations better resolving topography, or to LOCA's ability to capture the original GCM-predicted change in daily sequences of meteorology, which is not found in BCSD due to the use of historical analog months. In other cases, the comparisons may be complicated by differences in the VIC model parameters used in the two simulations. Changes in the VIC model parameters can result in changes in the climate change signal as expressed in hydrological variables [*Mendoza et al.*, 2015], and not changing the model parameters could result in a model calibrated for the statistics of the Maurer dataset, while being driven by meteorology with the statistics of the Livneh dataset. In LOCA-VIC, the parameter set is extremely similar to that used in BCSD-VIC, with minor adjustments made by the Livneh parameter set, but not a complete basin by basin recalibration (Appendix F). The best approach to take is an area of active research. It is not clear that either of these change signals should be considered more accurate, and the variability between GCMs remains larger than the variability between downscaling methods.

3.2.3. Explanation of differences from seasonal precipitation

Precipitation changes by the end of the century have a seasonal signature in CMIP5 GCMs. GCMs show increases across much of the northern part of the CONUS in boreal winter (Dec-Jan-Feb; DJF) and spring (Mar-Apr-May; MAM), and decreases in summer (Jun-Jul-Aug; JJA) across the Midwest and Pacific Northwest, as shown in the left column of Figure 17 for the RCP 8.5 scenario, with more modest changes in autumn (Sep-Oct-Nov; SON). The seasonal changes are larger than the yearly averaged results (Figure 14) because of partial cancellation of the wet winter signal with the dry summer changes. The middle column of Figure 17 shows the difference between the BCSD downscaled result and the original CMIP5 result, while the right column shows the analogous difference for the LOCA downscaled result. BCSD shows an overall shift towards wetter projections in the western half of the CONUS, particularly in winter and spring. (Alterations of the CMIP5-projected changes are also seen in the extreme southwest in summer and autumn in both LOCA and BCSD, however those regions are extremely dry during that part of the year, so the estimated changes measured in percent are subject to large sampling variability.) The shift can be large compared to the original CMIP5 prediction of change; for example, in DJF the CMIP5 GCMs project about a 25% increase in precipitation in the upper Midwest, and the additional increase in BCSD is about 15%. This is also reflected in a greater change signal in BCSD vs. LOCA (e.g., HUC 9 above and HUC 10 in Appendix D).
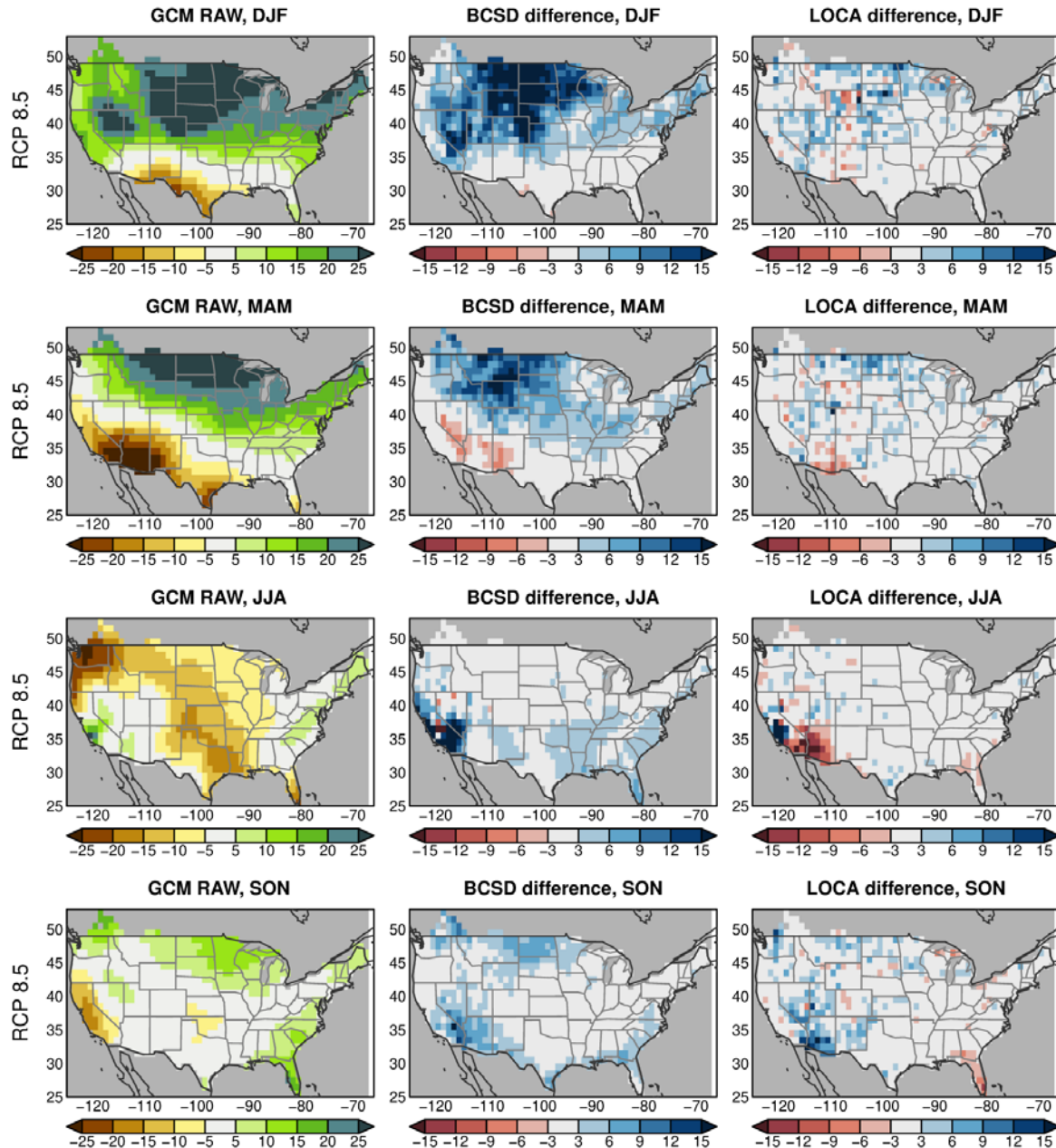
Figure 17 Direct comparisons with GCMs. Model-projected changes (%) in seasonal – Dec-Jan-Feb (DJF), Mar-Apr-May (MAM), Jun-Jul-Aug (JJA), Sep-Oct-Nov (SON) – precipitation, 2070-2099 with respect to 1970-1999, in the RCP 8.5 emissions scenario. Left column: multi-model ensemble averaged raw GCM results from the original 23 CMIP5 GCMs. Middle column: difference between the BCSD result and the CMIP5 GCMs. Right column: difference between the LOCA result and the CMIP5 result. All fields have been bilinearly interpolated to a common 1°x 1° grid for computing the difference.

3.3 Daily statistics

While most of this memorandum focuses on annual and seasonal time scales, differences in daily time scale statistics are also important to understand. LOCA was developed in part to more explicitly represent changes in extreme daily events, while BCSD inherits the monthly sequence of meteorology from the GCM and relies on historical weather sequences for daily features. Daily values in hydrology (and associated meteorology) are compared between LOCA-

VIC and BCSD-VIC to better understand how they represent extremes (annual maximum 3.3.1, flow duration 3.3.2, and how extremes change over time 3.3.3). Because daily meteorology can influence seasonal statistics of ET and runoff, we briefly present the percent of days where precipitation occurs (wet day fraction 3.3.4) and how it changes in future projections.

3.3.1 Annual maximum precipitation and runoff

To compare how extreme values differ between BCSD and LOCA, we compare the mean annual maximum precipitation (Figures 18, 19, 20) and runoff (Figures 22, 23, 24) for the historical period (1970-1999) (Figures 18, 22) and how these values differ in the future period (2070-2099) for RCP 4.5 (Figures 19, 23) and RCP 8.5 (Figures 20, 24).

As with previous statistics, the LOCA and BCSD datasets both mimic the statistics of the observation dataset they were trained with during the historical period. The LOCA dataset have lower mean annual maximum precipitation compared to BCSD in the historical period (Figure 18), the right column has more brown values than green particularly in California and the Canadian portion of the Columbia River basin. The pattern of differences between BCSD and LOCA across GCMs (different rows) are consistent with differences in the observational dataset (top row) indicating differences are largely driven by the observational dataset used in the downscaling, with the biggest differences being in California and areas in Canada. In addition to the differences between Livneh and Maurer, the BCSD dataset seems to exaggerate these differences slightly, with increases in the mean annual maximum precipitation compared to the Maurer dataset.

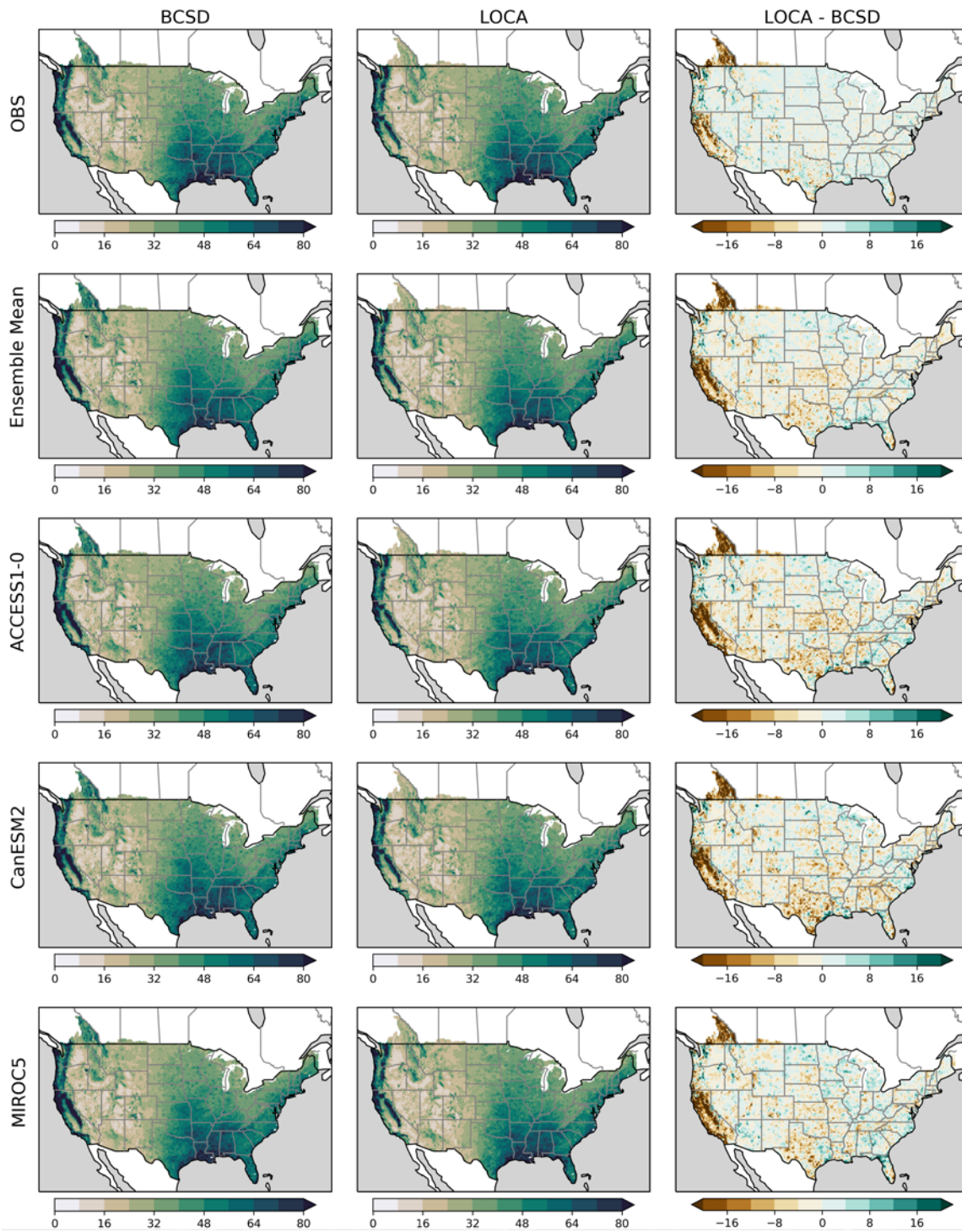# Precipitation (mm), Annual Maximum

Figure 18 Annual maximum precipitation differences. BCSD (left), LOCA (middle) and LOCA-BCSD (right) from observations (top row) averaged across all GCMs (row 2), and three individual GCMS (rows 3-5) averaged from 1970-1999.
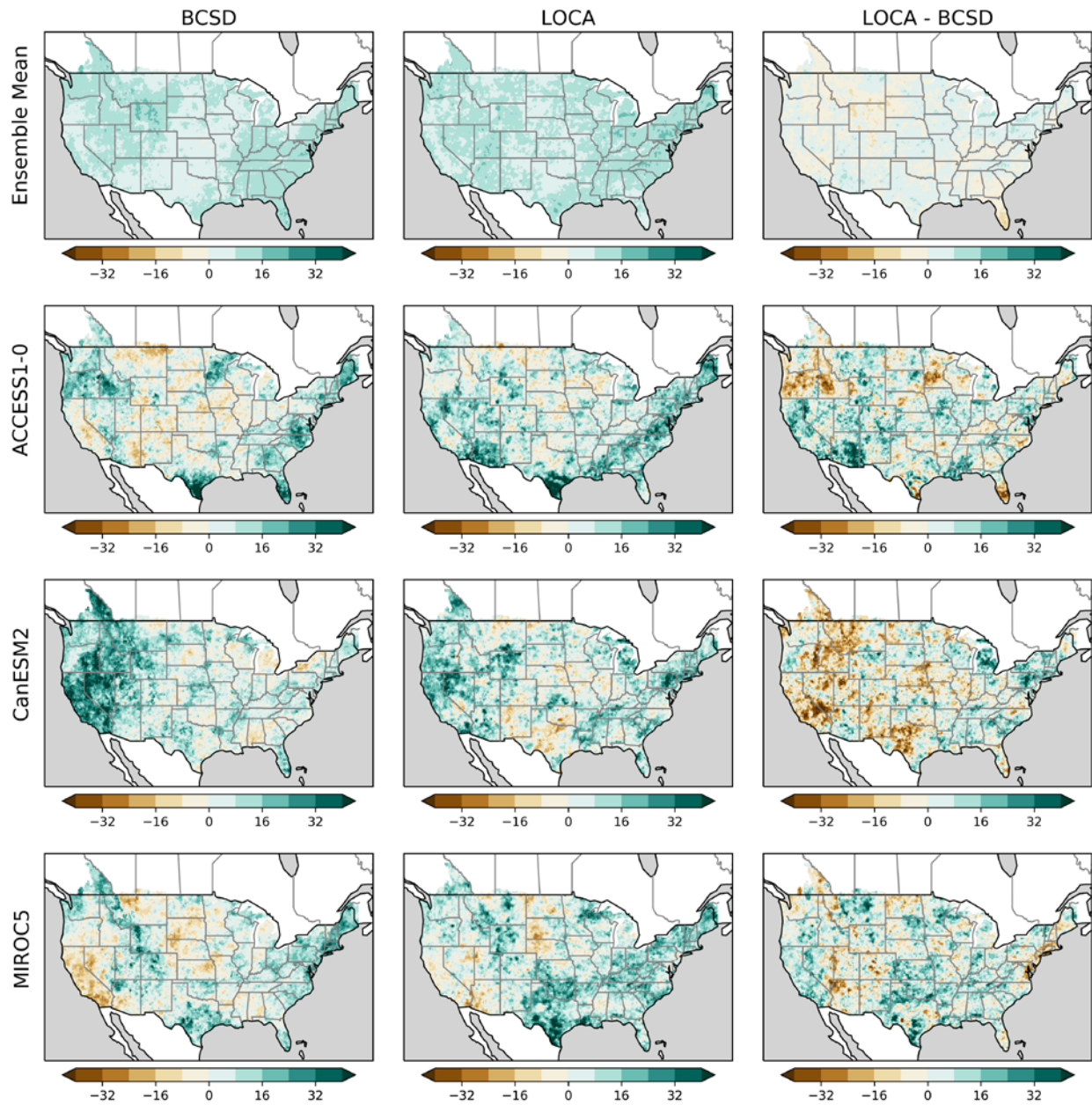
Figure 19 Changes in annual maximum precipitation. RCP 4.5. BCSD (left), LOCA (middle) and LOCA-BCSD (right). Differences between 1970-1999 and 2070-2099 averaged across all 23 GCMs (row 1), and three individual GCMs (rows 2-4).
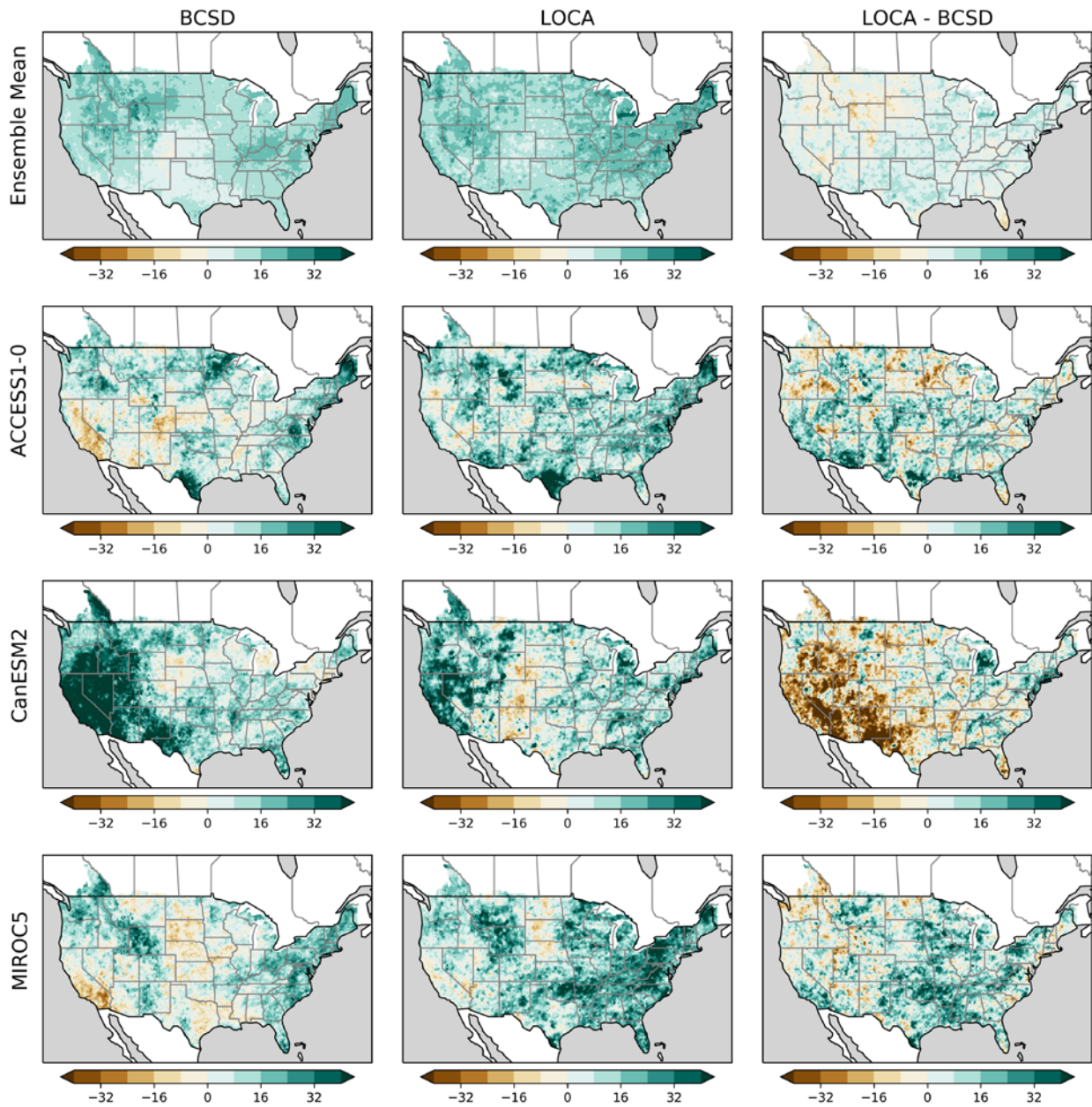
Figure 20 Change signals in annual maximum precipitation for RCP 8.5. Similar to the previous figure but for RCP 8.5 instead of 4.5.

In future projections, both methods show an increase in annual maximum precipitation in both RCP 4.5 (Figure 19) and RCP 8.5 (Figure 20) with increases in RCP 8.5 being greater than in RCP4.5. As with mean annual precipitation, the change in extreme precipitation in BCSD is associated with areas of higher topography (e.g. Figure 20, top left panel), while the changes in extreme precipitation in LOCA are more widely distributed. The differences between ensemble mean values for LOCA and BCSD are smaller than the change signal in most of the domain, though they are larger locally, for example the southern great plains, Eastern Colorado, and New Mexico. In contrast, when looking at individual GCMs the differences are comparable to the change signal (Figures 19, 20 bottom three rows). The changes in BCSD are associated

with increases in mean precipitation (e.g. the southwestern CONUS in CanESM2), and while LOCA often has similar associations, it is not as tightly coupled with changes in the mean because it works directly with the daily GCM output.

Changes in future extreme precipitation in the LOCA and BCSD downscaled datasets are ultimately driven by changes projected in the CMIP5 GCMs. These fields are compared in Figure 21. The left column shows the mean projected change (%) in 99th percentile precipitation in the 23 CMIP5 GCMs being downscaled, for the cold (Nov through Apr) and warm (May through Oct) seasons separately, for the RCP 8.5 scenario. The middle column shows the difference (percentage points) between the CMIP5 result and the analogous BCSD result, while the right column shows the difference for LOCA. Only days when precipitation is >= 1 mm/day are included in the analysis. The original CMIP5 data shows widespread increases in 99th percentile precipitation across most of the CONUS, with particularly large increases in the upper Midwest in the cold season and California during the warm season. Both LOCA and BCSD reproduce this pattern but with some modifications, with BCSD showing a substantially smaller change in California during the warm season than the CMIP5 models, and, to a lesser extent, the Southeastern seaboard and extreme Southwest during the cold season. LOCA has a reduced change compared to CMIP5 in Florida. The RCP 4.5 data show similar spatial patterns, although with reduced amplitude compared to RCP 8.5.
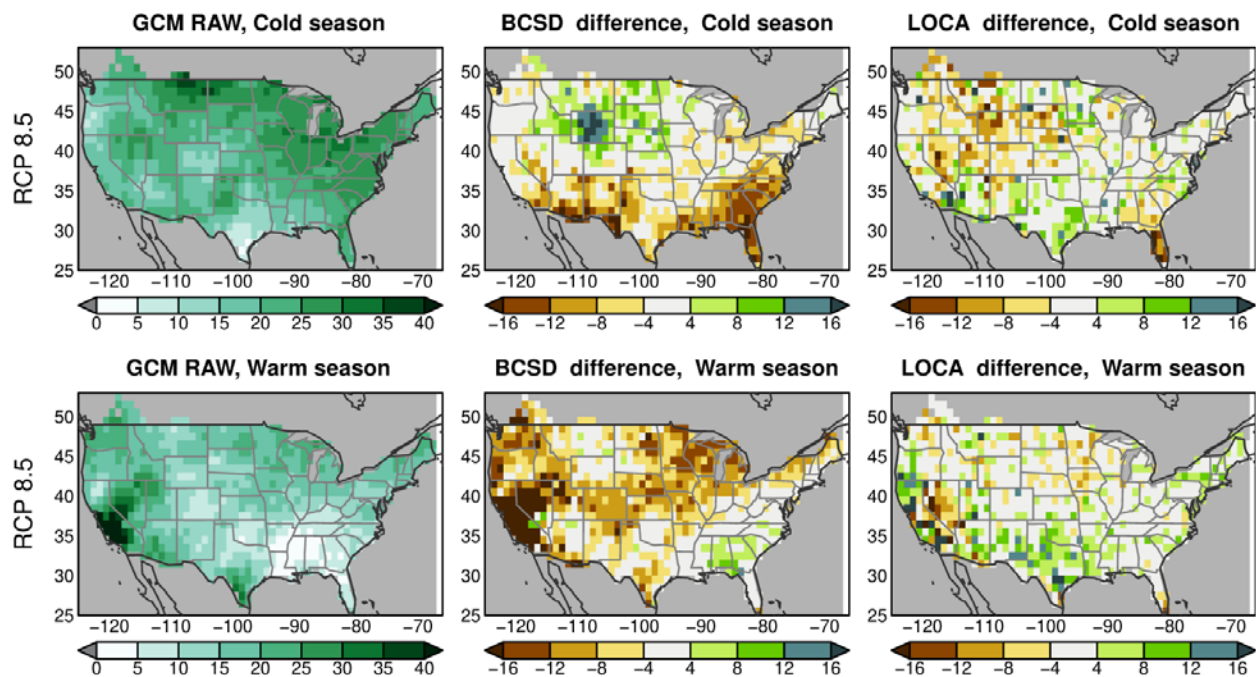

Figure 21 Ensemble averaged projected change (%) in 99th percentile precipitation. Values in the CMIP5 GCMs, 2070-2099 with respect to 1970-1999, for the RCP 8.5 scenario (left column). The upper row shows results for the cold season (NDJFMA), the lower row shows results for the warm season (MJJASO). Difference (percentage points) between the analogous BCSD projected change (middle column) and LOCA (right column) in 99th percentile precipitation with respect to the CMIP5 raw GCM result. All fields have been bilinearly interpolated to a common 1°x 1° grid for computing the difference.
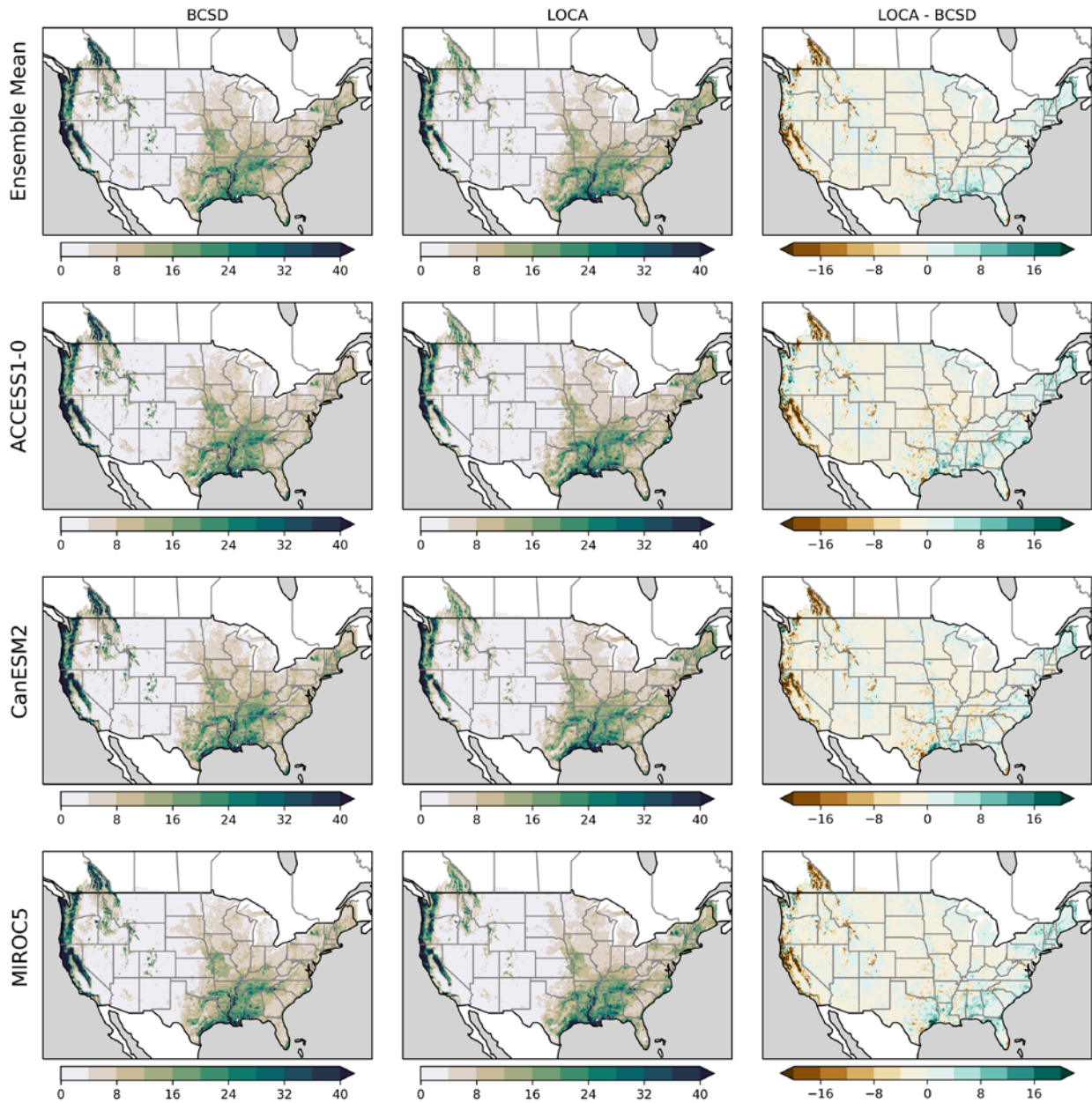
Runoff (mm), Annual Maximum

Figure 22 Annual maximum runoff (mm) differences. BCSD (left), LOCA (middle) and LOCA-BCSD (right) averaged across all GCMs (row 1) and three individual GCMS (rows 2-4). Only monthly runoff was available for Maurer so observations are not shown. All values are averages from 1970-1999.

The differences in mean annual maximum runoff generally follow the same patterns as mean annual maximum precipitation (Figure 22). In particular, the mean annual maximum runoff is generally similar, and the areas in which there are substantial differences, California and the Canadian portion of the Columbia River basin are the same as the BCSD and LOCA precipitation datasets. However, there are more consistently lower maximum runoff amounts in the western mountains in the LOCA-VIC dataset in the historical period compared to BCSD-VIC (Figure 22). This is likely related to the differences in mean annual precipitation (Figure 5) and the hydrograph in snowmelt dominated regions (Figure 16 HUCs 14 and 17), which have less

peak seasonal streamflow in LOCA-VIC than in BCSD-VIC. This may also be related to the higher elevation, colder snowpack in LOCA-VIC and delays in snowmelt occurring in the model which spreads the snowmelt out over a longer time period thus decreasing peak flows, and delays it to time periods when potential ET is greater and vegetation is more actively transpiring. In addition, the land surface model processes may play a more important role in modulating runoff in the central CONUS, with sharp spatial gradients evident in, e.g., North Central Kentucky, or Western Iowa (Figure 22) that are not present in the mean annual maximum precipitation patterns (Figure 18), and are most likely driven by hydrologic processes and spatial variations in model parameters.
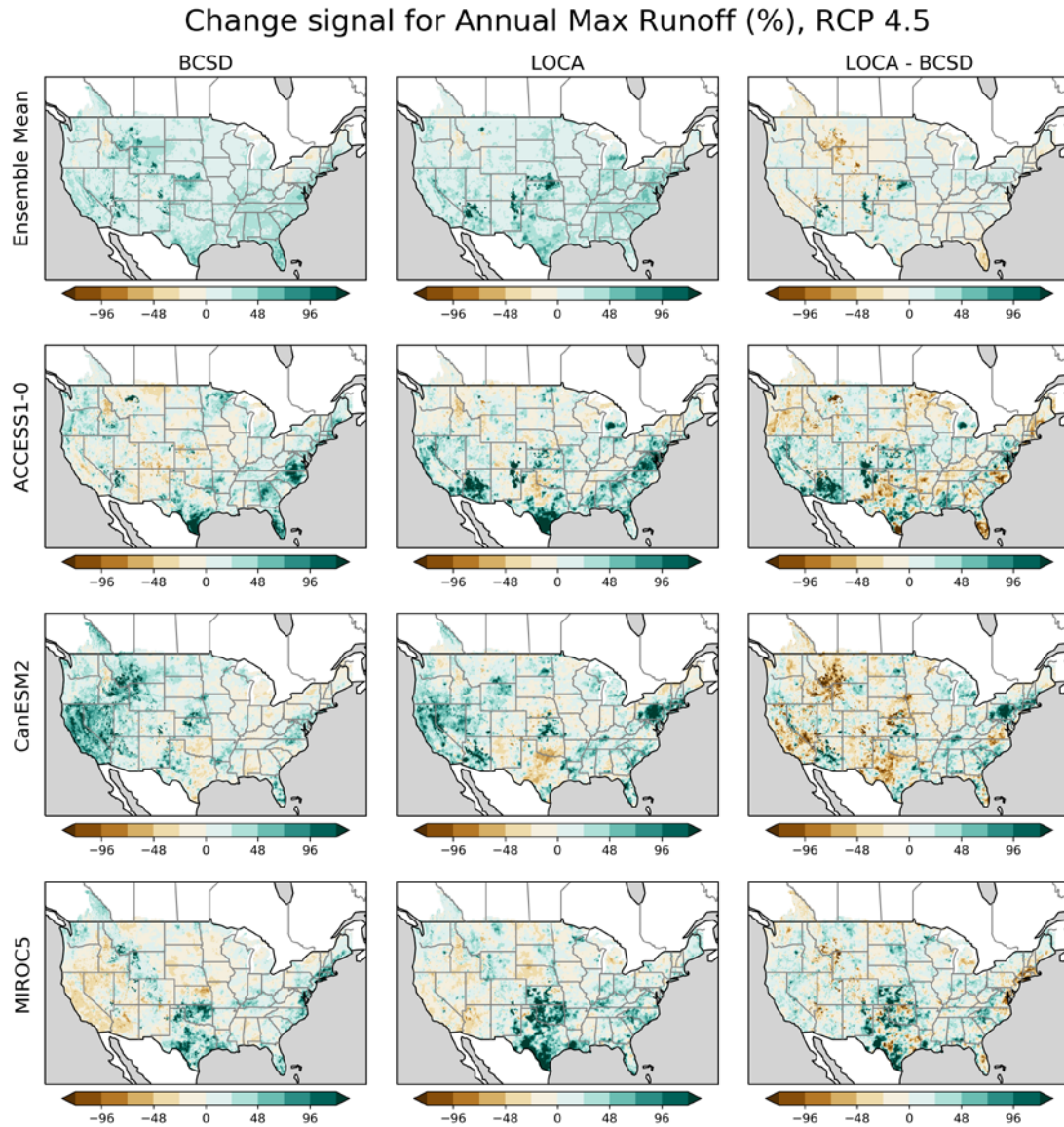
Figure 23 Change in annual maximum runoff for RCP 4.5. Similar to Figure 19, but for annual maximum runoff, instead of annual maximum precipitation.

The changes in mean annual maximum runoff in LOCA-VIC and BCSD-VIC (Figure 23, 24) closely follow the changes in mean annual maximum precipitation in LOCA and BCSD (Figures 19, 20), with some sharp spatial gradients caused by the VIC land surface properties evident in

the ensemble mean. As with precipitation, both methods show an increase in mean annual maximum runoff in both RCP 4.5 (Figure 23) and RCP 8.5 (Figure 24, with changes in RCP8.5 being greater than changes in RCP 4.5. The spatial patterns of changes in VIC runoff projected with individual GCMs (bottom three rows) closely follow the spatial pattern of changes in precipitation from those individual GCMs. The ensemble mean changes (top row), however, show larger than average increases in a few regions in the central and western CONUS in both LOCA-VIC and BCSD-VIC, for example in Southern Nebraska, Northern Kansas, Western Arizona. These spatial patterns are most likely linked to the VIC hydrologic model parameters in these regions, as there are very few changes in extreme precipitation here. These features are more pronounced in the LOCA-VIC dataset in the central CONUS, and in the BCSD-VIC dataset in the northern interior mountains. This is due to the combination of model parameters and changes in extreme precipitation in the two datasets. Model parameters are likely responsible for the sharp features in the central CONUS, where BCSD has relatively little change in extreme precipitation compared to LOCA (Figures 23, 24). The changes in the northern interior mountains may be due more to the sharp gradients in changes in extreme precipitation in BCSD in these regions compared to LOCA. Both features need to be investigated further to understand the causes of the changes in extreme runoff before any confidence can be placed in them.
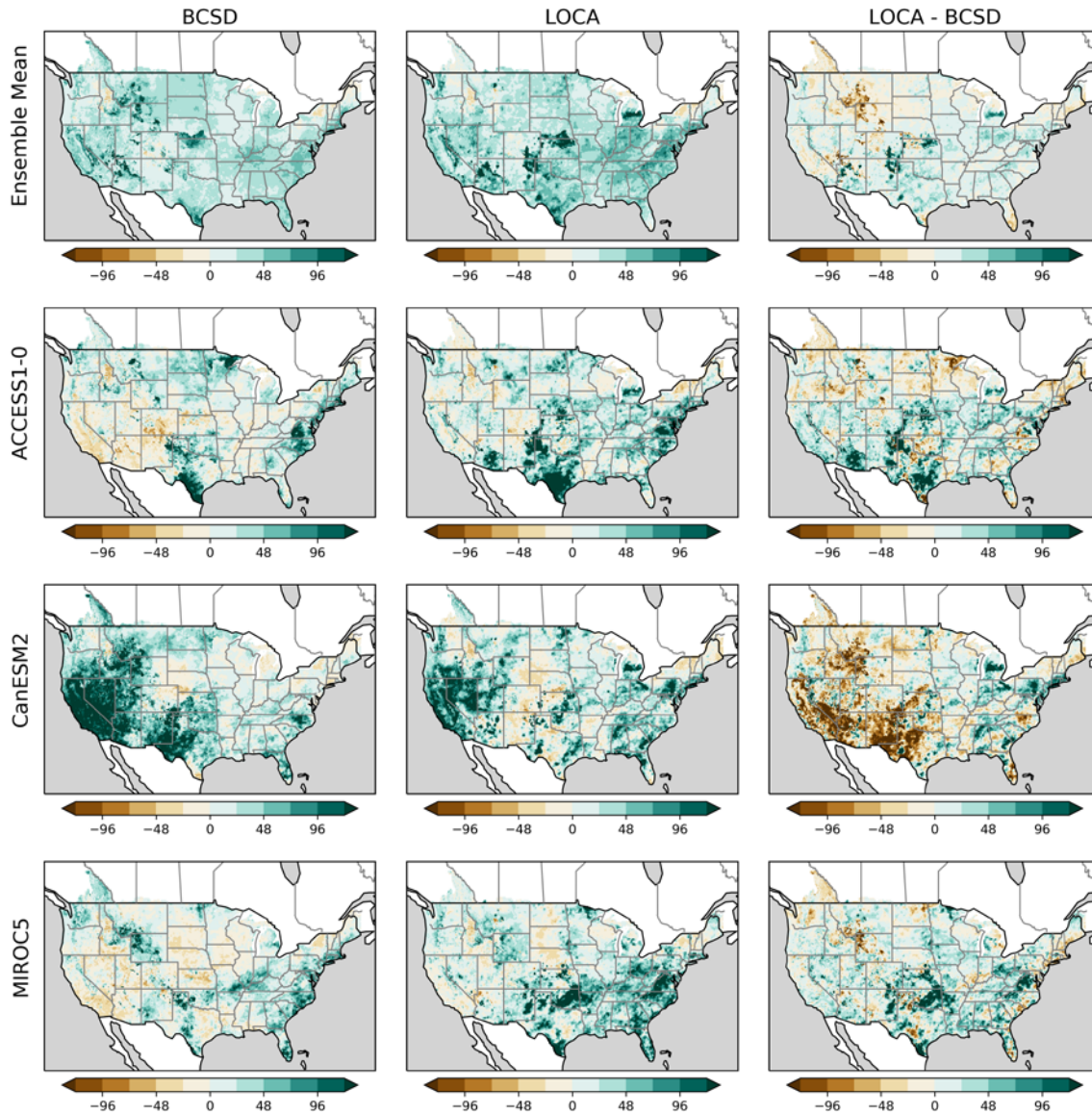
Change signal for Annual Max Runoff (%), RCP 8.5

Figure 24 Change in annual maximum runoff for RCP 8.5. Similar to the previous figure but for RCP 8.5 instead of 4.5.

3.3.2. Flow duration

To better understand streamflow differences, especially daily high and low values, we compare the daily flow duration curves based on routed runoff from LOCA-VIC and BCSD-VIC in several western basins. These daily flow values have not been bias corrected and these basins are not consistently calibrated. Some basins have likely been calibrated by one researcher or another over the years, but no detailed catalog of the calibrations exists. In addition, different researchers have used different calibration strategies, so the calibrations that have been performed are not consistent, and in some regions no calibration has been performed at all. Therefore, the modeled flows should not be expected to capture historical magnitudes accurately (careful considerations would need to be made before using simulations in operations models). The routing is done with a network routing scheme [*Mizukami et al.*, 2016]

that uses the USGS geospatial fabric [*Viger*, 2014]. This provides an identical network setup and routing algorithm and thus allows a direct comparison between the two datasets so that differences are attributable to the differences in forcing datasets and hydrologic model configuration. Even though the routed flows to not match observations, these values give us a general sense of how the two datasets differ by looking at relative differences. We present results for 43 basins in the CONUS (Figure 25, Table 2) for the historical (1970-1999) and future (2070-2099) for RCP 4.5 and 8.5. These basins were selected as they are the same locations identified in *Reclamation* [2014].

Figure 25 Basins evaluated for flow duration. Figure from Reclamation [2014], see Table 2 for basin legend.

The daily flow duration curves for seven of the 43 basins are shown in Figure 26. Generally, flow duration curves are similar between BCSD-VIC and LOCA-VIC (Figure 26, see Appendix E for all 43) with the 23-GCM ensemble spread being wider in the future than in the historical period for both downscaling methods. However, the flow duration curves are not identical, and how they differ varies between basins. In some locations they are nearly identical (e.g., Yakima at Parker, Colorado at Lees Ferry); in some locations LOCA flows are greater (e.g., Williamson below Sprague, Feather at Oroville); in some locations BSCD flows are greater (e.g., Columbia at Grand Coulee, Plate South Fork near Sterling); and at times (though less common) their paths cross (e.g, in the Rio Grande near Lobatos, LOCA has higher high flows and lower low flows than BCSD), indicating a much flashier response of streamflows. We do not intend these basins to be a comprehensive set, but rather specific locations of interest based on past reports. We therefore do not make overarching conclusions but rather display the information so those interested can look more closely at basins they are interested in and contrast them to others.

Notably too, these are modeled values that often have biases in excess of their change signal in both BCSD-VIC and LOCA-VIC. We intend these figures to illustrate differences from downscaling methods prior to any streamflow bias correction. If streamflow values were to be used beyond relative comparisons, closer investigations would be required, and likely biases would need to be corrected.

Table 2 Basins Evaluated for Flow Durations. Locations similar to *Reclamation* [2014].

| Number	State	River Basin and Outlet Location	Latitude	Longitude
1	OR	Williamson River below Sprague River	42.56	-121.84
2	CA	Klamath River below Iron Gate Dam	41.93	-122.44
3	CA	Klamath River below Seiad Valley	41.85	-123.23
4	CA	Klamath River at Orleans	41.30	-123.53
5	CA	Klamath River near Klamath	41.51	-123.98
6	ID	Snake River at Brownlee Dam	44.84	-116.90
7	WA	Columbia River at Grand Coulee	47.97	-118.98
8	OR	Columbia River at the Dalles	45.61	-121.17
9	WA	Yakima River at Parker	46.51	-120.45
10	OR	Deschutes River near Madras	44.73	-121.25
11	ID	Snake River near Heise	43.61	-111.66
12	MT	Flathead River at Columbia Falls	48.36	-114.18
13	AZ	Colorado River at Lees Ferry	36.86	-111.59
14	CA	Colorado River above Imperial Dam	32.88	-114.47
15	UT	Green River near Greendale	40.91	-109.42
16	CO	Colorado River near Cameo	39.24	-108.27
17	CO	Gunnison River near Grand Junction	38.98	-108.46
18	UT	San Juan River near Bluff	37.15	-109.86
19	CA	Sacramento River at Freeport	38.46	-121.50
20	CA	Sacramento River at Bend Bridge (Red Bluff)	40.26	-122.22
21	CA	Feather River at Oroville	39.52	-121.55
22	CA	San Joaquin River near Vernalis	37.68	-121.27
23	CA	Stanislaus River at New Melones Dam	37.95	-120.53
24	MT	Missouri River at Canyon Ferry Dam	46.65	-111.73
25	MT	Milk River at Nashua	48.13	-106.36
26	CO	Platte River (South Fork) near Sterling	40.62	-103.19
27	NE	Missouri River near Omaha	41.26	-95.92
28	CO	Rio Grande near Lobatos	37.08	-105.76
29	NM	Rio Chama near Abiquiu	36.32	-106.60
30	NM	Rio Grande near Otowi	35.88	-106.14
31	NM	Rio Grande at Elephant Butte Dam	33.16	-107.19
32	NM	Pecos River at Damsite No 3 (Carlsbad)	32.51	-104.33
33	CA	Little Truckee River below Boca Dam	39.39	-120.10
34	CA	Carson River (West Fork) at Woodfords	38.77	-119.83
35	CA	Sacramento-San Joaquin Delta inflow	38.06	-121.86
36	CA	San Joaquin River at Friant Dam	37.00	-119.71
37	CA	Truckee River at Farad Gage (stateline)	39.45	-120.01
38	NV	Truckee River at Nixon Gage	39.78	-119.34
39	NV	Carson River at Ft Churchill Gage	39.33	-119.15
40	MT	Big Horn River at Yellowtail Dam	45.31	-107.96
41	NE	Platte River (North Fork) at Lake McConaughy	41.21	-101.64
42	CA	American River at Fair Oaks	38.64	-121.23
43	CA	Tulare-Buena Vista Lakes basin	36.05	-119.72
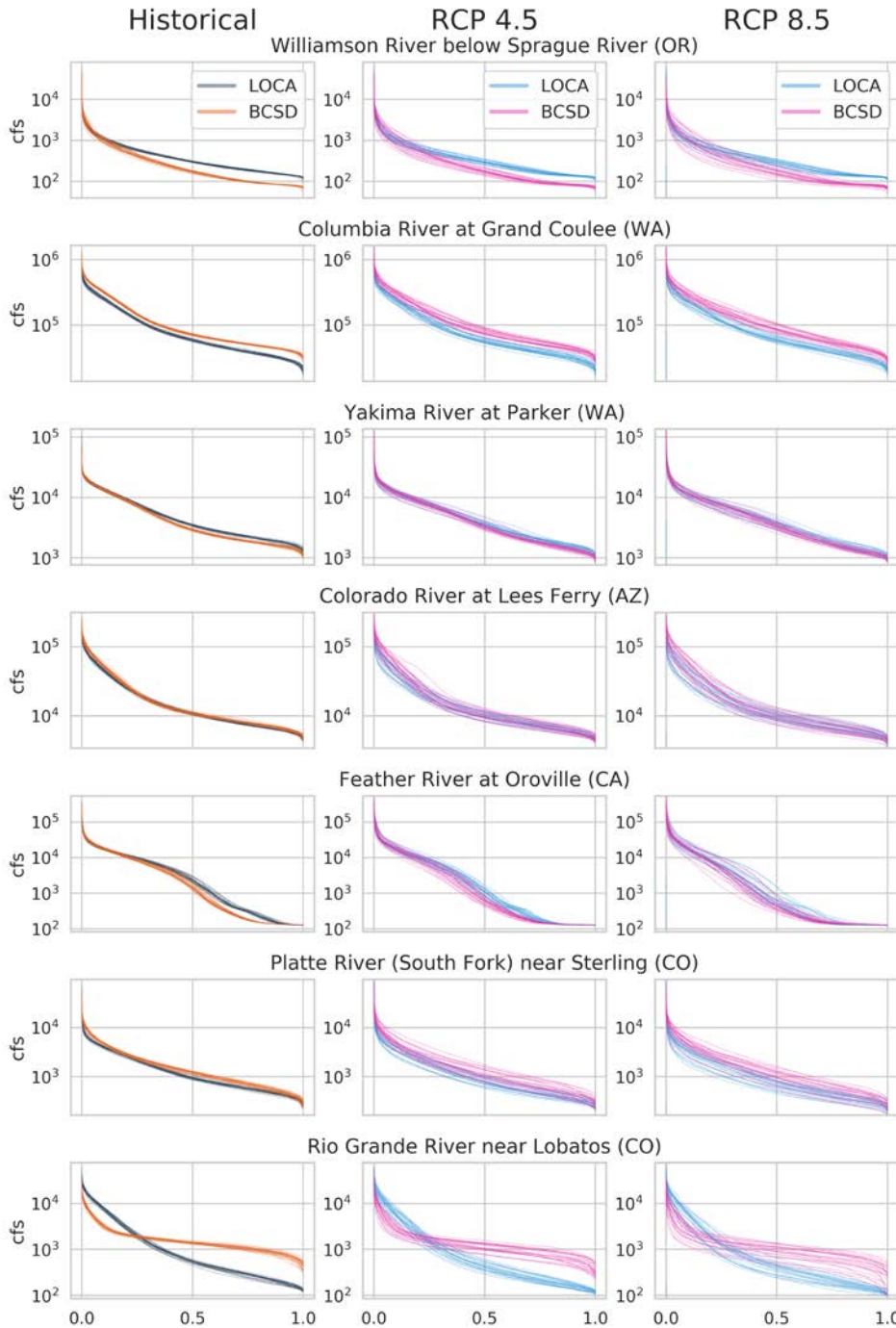
Figure 26 Daily Streamflow Exceedance Probability. Historical (left column) and RCP 4.5 (middle) and RCP 8.5 (right). Individual lines come from routed flows from BCSD-VIC or LOCA-VIC run with 23 GCMs. These are seven of the 43 basins in Table 2, other basins can be found in Appendix E.

3.3.3 Extreme runoff changes over time

Both LOCA and BCSD project an increase in the frequency of extreme daily runoff events in the future. Figure 27 shows box plots of the number of grid cells in each 5-year period across the

entire domain for which the downscaled hydrology dataset has an annual maximum runoff event greater than or equal to the largest daily event in that grid cell in the historical period. A step change in the frequency of extreme precipitation and runoff events also appears in both LOCA and BCSD when transitioning from the historical period to the future projection period (2005 onward for LOCA, 2000 for BCSD). The step change produces a greater number of extreme events in the future period. The change in the seventy-fifth percentile is greater in BCSD, while the change in the median in the LOCA dataset is larger. This step change most likely comes from constraints applied in the historical period such that the GCM precipitation is forced to match the observed variability, while in the projection time period, these constraints are removed.
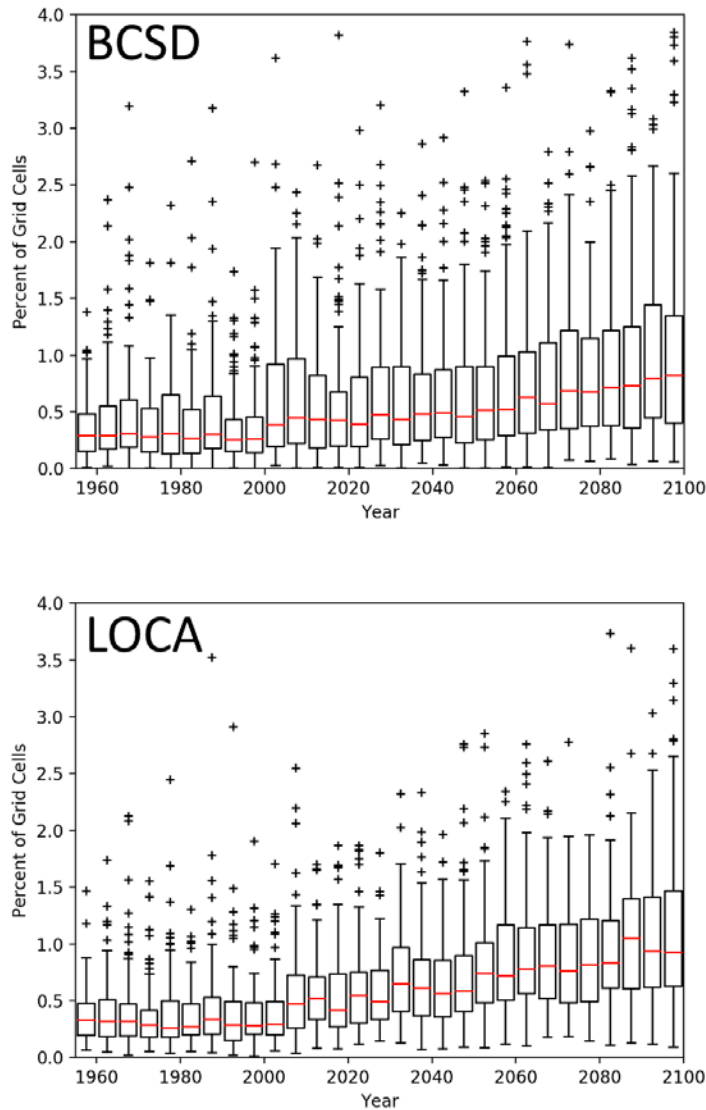
Figure 27 Trends in frequency of extreme runoff events for RCP 8.5. Box plots of the number of grid cells in each five-year period for which the historical maximum runoff amount was equaled or exceeded.
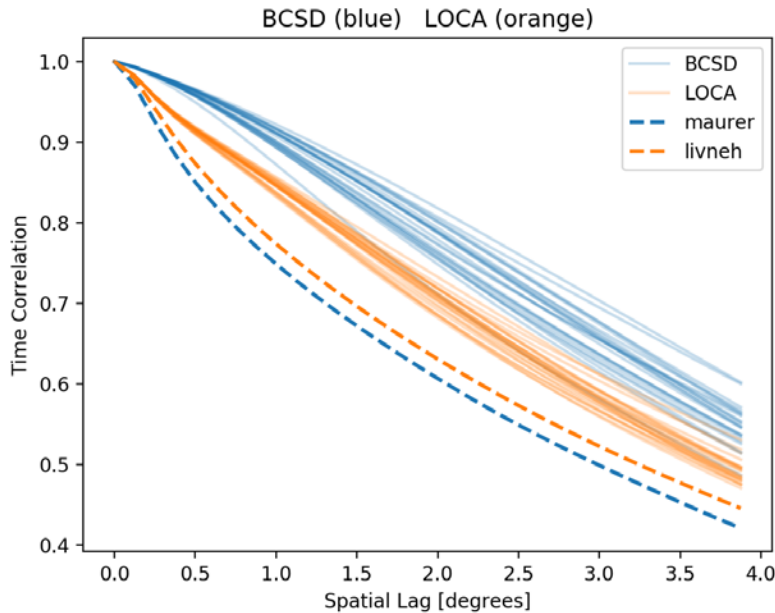
Figure 28 Historical space-time autocorrelations. The monthly time series correlation coefficients between all pairs of grid cells with a given spatial separation (lag) as averaged across the domain for BCSD (all GCMs, light blue), Maurer (blue), LOCA (all GCMs, light orange) and Livneh (orange) for precipitation.

Another difference between LOCA and BCSD downscaling methods is evident in the space-time autocorrelation between grid cells. To illustrate this, the correlation in time between each grid cell with nearby grid cells of different proximities is computed for the historical period. The correlation values across grid cells are grouped by distance, averaged, and plotted for each GCM and the observations in Figure 28. This is important to hydrological applications because an unrealistically high spatial coherence of the precipitation field can skew flooding or regional drought statistics. For example, if adjacent watersheds are more likely to receive precipitation simultaneously in a downscaled dataset than in the observations, anomalously large floods may be simulated at their confluence. Both downscaled datasets produce realistic space-time autocorrelations on the daily time scale; however, differences are evident on the monthly timescale. Because BCSD uses the GCM precipitation on a monthly time scale with a quantile-mapping bias correction and a simple disaggregation routine, it inherits more of the GCM spatial autocorrelation distance in its precipitation field. In contrast, LOCA operates on a daily time scale and predicts a different high-resolution field based on the spatial pattern in the GCM on every day. As a result, LOCA produces more spatial variability on a monthly time scale than was originally projected by the GCM. This may lead to more spatially coherent droughts or floods in BCSD as compared to the LOCA. It is also clear that neither method reproduces the spatial variability that appears in the observations, though LOCA is less affected than BCSD. This reinforces the point that it is important to perform any climate change analysis based on a consistent dataset in the historical period (e.g. LOCA future – LOCA present, not LOCA future – observed).

3.3.4 Wet day fractions

Wet day fraction, the fraction of days in which there is precipitation, is a particularly important value in hydrologic modeling as it strongly influences ET. Forcing data with positive biases in wet day fraction often have too much drizzle and, as a result, too much ET particularly from

vegetation canopies. However, when running a hydrologic model it is common to estimate a daily cycle of temperature, humidity, and radiation based on the diurnal temperature range and the presence or absence of precipitation [*Bohn et al.,* 2013]. In particular, the VIC hydrologic model uses the MTCLIM [*Thornton and Running,* 1999] algorithm to estimate these forcing data (in VIC versions used here, this is done internally, in newer versions this is done as a pre-processing step). MTCLIM assumes that when precipitation is present clouds are too, and it decreases the amount of shortwave radiation in the model, which has the effect of decreasing ET. As a result, the effect of wet day fraction on ET is not straightforward to diagnose.

In the historical period, the observed datasets (top row, Figure 29) Maurer (for BCSD) and Livneh (for LOCA) have similar spatial patterns in wet day fractions. Livneh has more wet days than Maurer, especially in Canada and California. Although the spatial patterns in LOCA and BCSD are broadly similar, this is a case in which the historical GCM ensemble means do not have similar differences between BCSD and LOCA as the observed datasets used in their downscaling, with LOCA having a lower wet day fraction than BCSD. GCMs tend to have an overly high wet day fraction (the so-called "drizzle problem"), so LOCA directly bias-corrects the wet day fraction, resulting in values closer to the observed dataset as seen in Figure 29. The variability across GCMs in BCSD matches the variability in precipitation across GCMs, and is greatest in and around California-Arizona, while it matches the variability in wet day fraction across GCMs (not shown) in LOCA in Texas-Louisiana.   Both of these are explained by the methods used in the two downscaling methods. LOCA explicitly represents daily variability in precipitation based on the GCM patterns, while BCSD uses monthly sequences of wet days from the observations. However, BCSD also adds additional wet days when the extreme precipitation event exceeds the maximum value in the historical period. The excess precipitation in BCSD is distributed across all days in the month. As a result, greater variability in monthly precipitation totals can cause BCSD to increase the number of wet days.

In future projections (Figure 30 and 31), wet day fraction from BCSD has little change because it does not inherit the daily sequence from the GCMs, whereas the fraction in LOCA changes dependent on the GCM changes. For LOCA, these changes are larger in the higher emissions experiments. On average, LOCA's wet day fractions decrease in the future, although this differs across GCMs.

In these comparisons, note that the LOCA data has been aggregated to match BCSD's 1/8th degree grid. This leads to more wet days at any particular location than what was actually used to drive the hydrologic model.
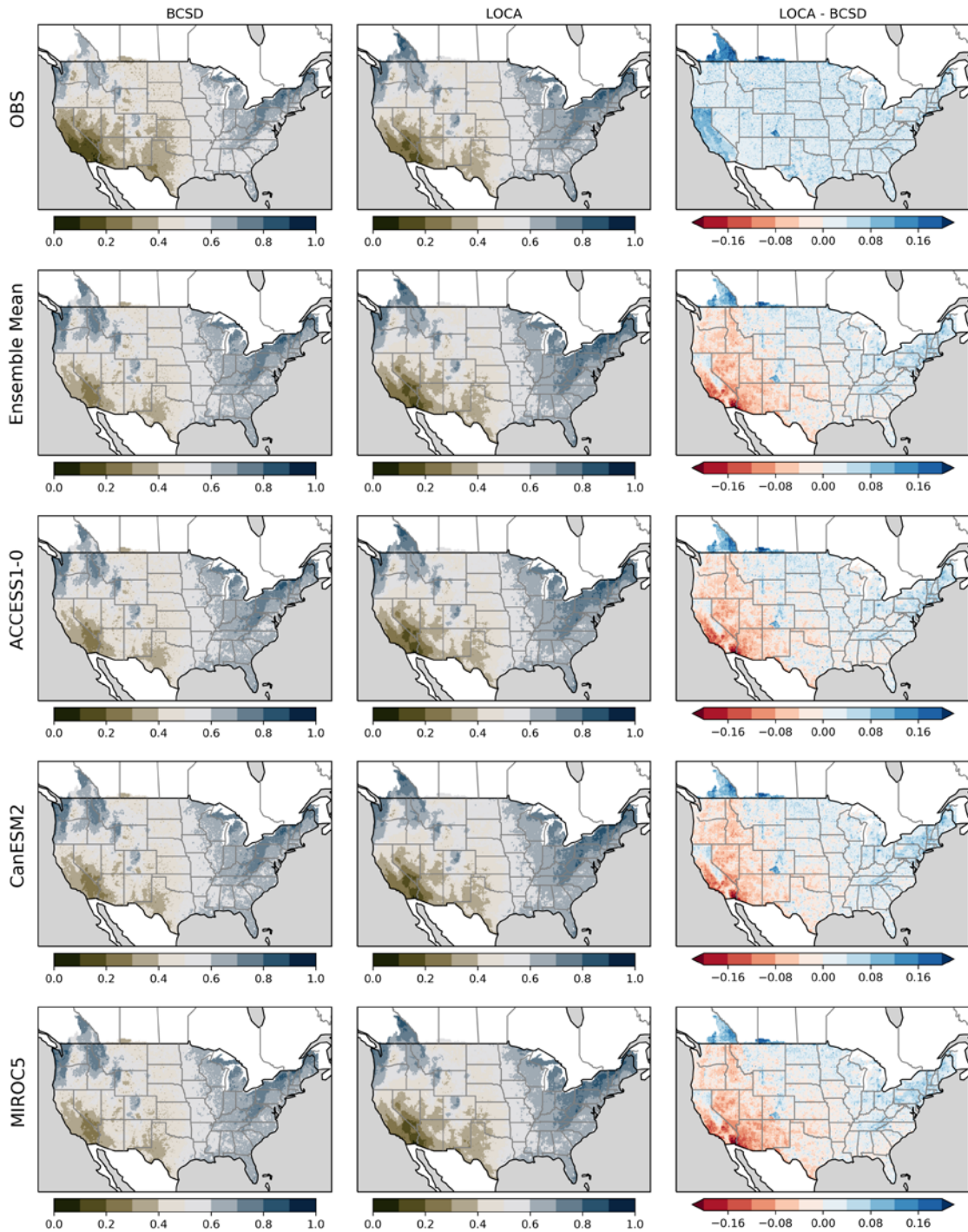
Precipitation (mm), Wet Day Fraction

Figure 29 Historical Wet Day Fraction. Maps of the fraction of days for which precipitation is present (non-zero) from BCSD (left), LOCA (middle) and LOCA-BCSD (right) from observations from Maurer (left) and Livneh (middle) (top row) averaged across all GCMs (row 2), and three individual GCMS (rows 3-5) averaged from 1970-1999.
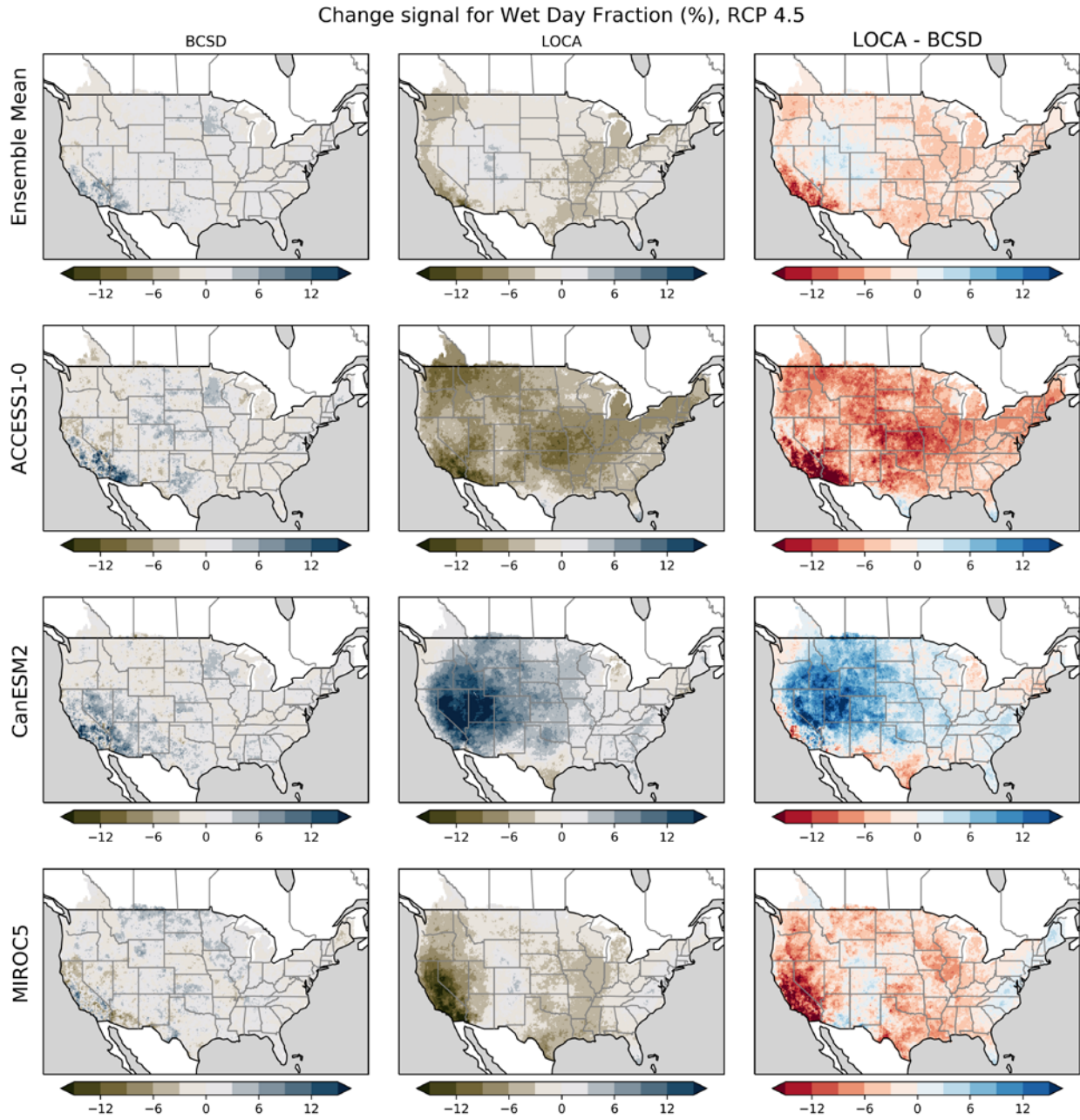
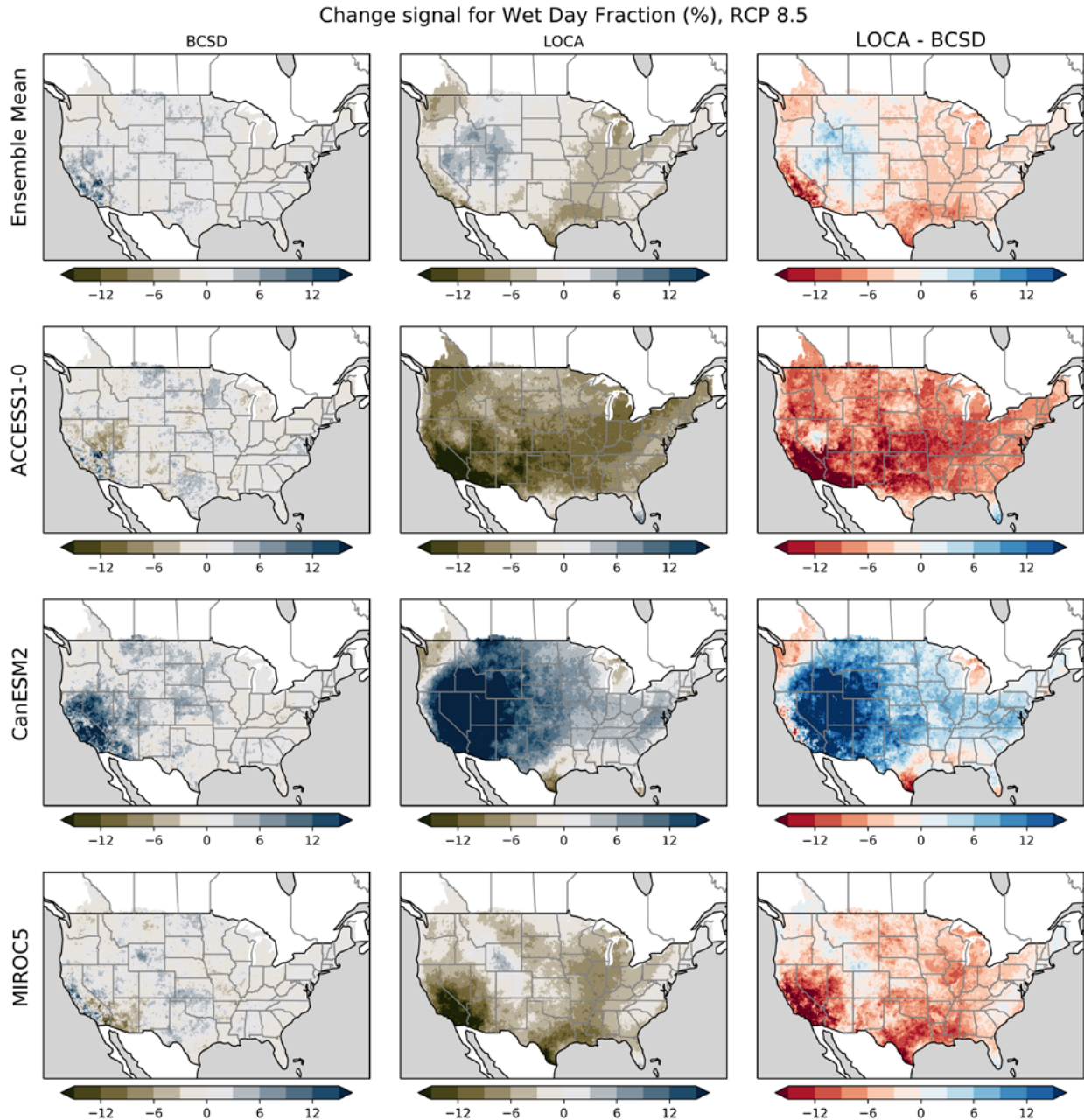**Figure 30 Change in Wet Day Fraction for RCP 4.5.** Similar to Figure 19, but for wet day fraction, instead of annual maximum precipitation.

Figure 31 Change in Wet Day Fraction for RCP 8.5. Similar to the previous figure but for RCP 8.5 instead of 4.5.

The projected changes in wet day fraction shown in Figure 30 and 31 correspond to projected changes in the original GCMs, which show a spatial signature with 3-15 fewer wet days/year in the Pacific Northwest, Southwest, and southern parts of Texas (Figure 32, left column), depending on the location and emissions scenario. This is computed with a wet day threshold of >= 1 mm/day, necessary since the low-resolution GCMs produce too much drizzle. Note that this is a different threshold than used in Figures 29-31 (which have a zero threshold to demonstrate MTCLIM sensitivities, described above). Compared to the CMIP5 GCM historical-to-future changes, LOCA tends to project a change of 3-6 more wet days through Utah, Nevada, Idaho, Eastern Oregon, and Eastern Washington, and 3-6 fewer wet days/year in much of the

East. BCSD has a more monotonic pattern of increasing the number of wet days by at least 3/year days over the entire domain, with the largest shifts with respect to the CMIP5 result in the western CONUS, particularly the Pacific Northwest.
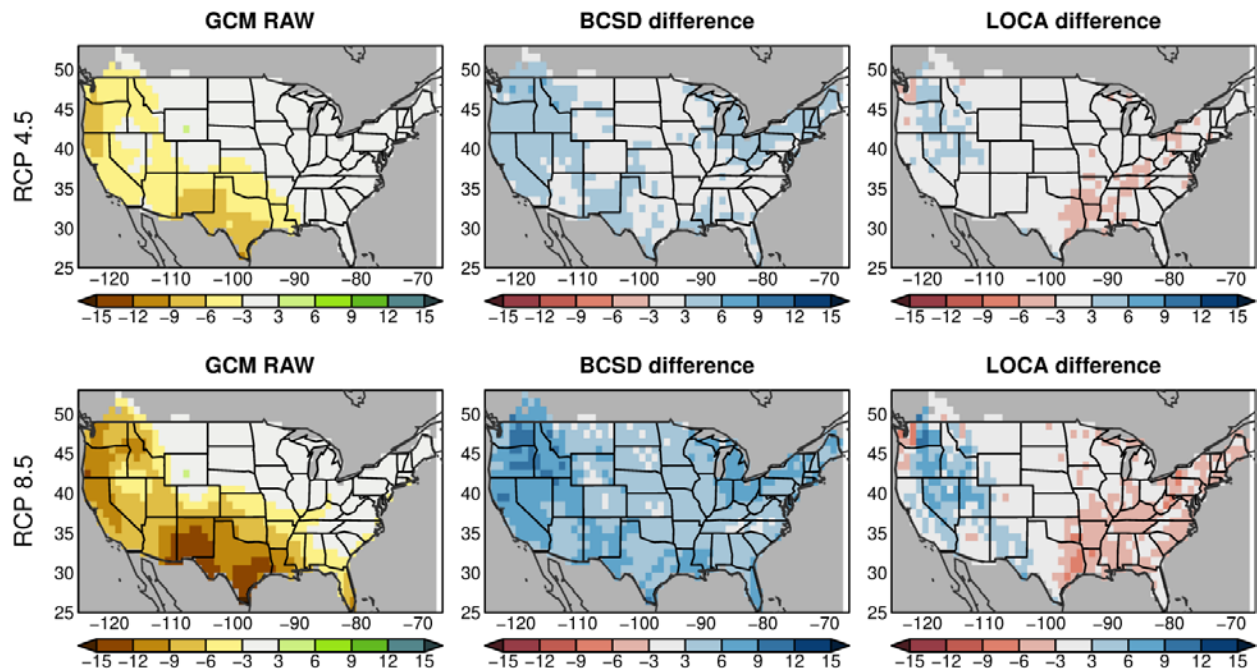
Figure 32 Direct GCM comparison of wet days. Left column: projected change (days) in number of wet days per year, 2070-2099 with respect to 1970-1999, for RCP 4.5 (top row) and RCP 8.5 (bottom row) emissions scenarios. Left column: Change projected by the CMIP5 GCMs. Middle column: difference between the BCSD result and the CMIP5 GCMs. Right column: difference between the LOCA result and the CMIP5 result. All fields have been bilinearly interpolated to a common 1°x 1° grid for computing the difference. In this figure, a wet day is defined as P >= 1 mm/day.

3.4 Other notable differences

In addition to differences between the LOCA and BCSD products noted above, there are differences that occur because of decisions made during the process of generating the data.   In this evaluation, we uncovered two worth noting:

1. Four models, CCSM4, GISS-E2-R, CESM1-CAM5, and FGOALS-g2, used different historical period simulations in BCSD vs. LOCA. LOCA used ensemble member r6i1p1 for CCSM4 and GISS-E2-R because when LOCA was undertaken, daily data was not available for ensemble member r1i1p1, which is what BCSD used. For CESM1-CAM5 and FGOALS-g2, the difference between LOCA and BCSD is likely due to a reposting of the GCM data between when the BCSD and the LOCA downscaling efforts downloaded their raw GCM data. This difference was detected by doing a simple correlation of annual precipitation anomalies from BCSD and LOCA over the Upper Colorado River (HUC4). Because BCSD and LOCA both derive their interannual variability from the GCM, timeseries of mean annual precipitation anomalies are expected to be well correlated. More detail on these differences is available on the CMIP5 Errata page: https://cmip.llnl.gov/cmip5/errata/changes_of_FGOALS-g2_historical_data.pdf
2. Differences exist between methods depending on how downscaling regions were defined. For BCSD, downscaling was performed by HUC2 and later merged into the full

CONUS domain. In some cases, this resulted in artifacts at the HUC boundaries (see Figure 4). The squiggly lines across the BCSD difference plots are related to the HUC2 boundaries, BCSD was implemented independently for each HUC2 and as a result there can be artifacts on the boundaries in a continental domain analysis. For LOCA, the 1x1 degree downscaling boxes are apparent for some measures, such as wet day fraction.

4. Results in Context and Conclusions

Statistically downscaled climate data has been at the core of hydroclimate projections for over a decade. This is a result of the need for meteorological data that are consistent with the requirements of hydrologic models so river basins and local watersheds can be modeled to provide a reasonable approximation of the real hydrologic system. Importantly, this approximation must be consistent in the historical and future configurations to provide plausible climate change information.

Through this evaluation, we assess that both LOCA and BCSD and their associated VIC hydrologic simulations provide reasonable reproductions of current climate and hydrology with some notable departures. Differences between LOCA-VIC and BCSD-VIC or between downscaling methods and observation-driven VIC simulations are most pronounced in mountainous environments. Climatological differences between LOCA-VIC and BCSD-VIC simulations are largely due to differences between the Livneh and Maurer observed datasets, this finding is consistent with other recent comparisions [*Alder and Hostetler*, 2019, *Jiang et al.,* 2018]. Differences are more pronounced in the mountains generally, and in California and the Canadian portion of the Columbia River basin in particular. These differences are more or less notable depending on the hydroclimate variable of interest. Technical differences also translate to differences in modeled hydrology including differences in resolution (e.g., 1/8 vs 1/16 degree), different domain boundaries, and the difference in several combinations of GCM/RCP, although these elements had lesser importance in explaining differences in this particular comparison as it has been limited to identical GCM/RCP combinations.

Both methods produce plausible future climate projections. However, it is important to recognize that they are not the same and any individual climate projection is a possible future not a specific prediction. Generally, we found the average monthly and annual changes in precipitation in the LOCA dataset are smaller than the average monthly and annual changes projected by BCSD. This results from the requirement in LOCA that the change signal in the GCM should be preserved, while BCSD allows the bias correction to modify the change signal (e.g. regions with more precipitation variability in the GCM than observed have the magnitude of the climate change signal reduced, while regions with less precipitation variability than observed have the magnitude of the climate change signal increased [*Maurer and Pierce,* 2014]). The extent to which change signals should be preserved is, however, an open research question (c.f. *Maraun* [2016] for a review of the issue).

Two important implications of these differences are: (1) when the LOCA data are used, it is important to recalibrate VIC (or other models) to be consistent with the LOCA climate statistics, (2) one should always compare changes within a consistent modeling framework. In other words, avoid direct comparisons between LOCA and BCSD derived flows and avoid direct comparisons between observations and future GCM simulations. Instead, compare simulations in a historical GCM to the same GCM in the future downscaled using the same method.

In addition, while BCSD and LOCA predict very similar change in long-term statistics, LOCA explicitly projects changes in many daily statistics based on the GCM daily weather sequences. As a result, for certain statistics and in certain locations, the two methods diverge. For example, the methods have considerable differences in the mountains for extreme precipitation and runoff. This, in part, has to do with BCSD being tied to changes in monthly statistics and climatological patterns, such as greater precipitation corresponding to more extreme precipitation in the mountains, while LOCA values are able to vary more. We also note that both LOCA and BCSD have a statistical artifact in the frequency of extreme precipitation and runoff

that manifests as a step change in the change in frequency between the historical and future periods, at year 2000 for BCSD and 2005 for LOCA. This is a statistical artifact, and users are encouraged to consider using the period 2001-2020 as the historical reference (BCSD) or 2005-2025 (LOCA) as in Wobus et al [2017]. Beyond extreme events, LOCA also projects changes in the number of wet days in the future, while BCSD exhibits only limited changes; such changes may affect ET and runoff, though changes in precipitation have a much greater effect.

Finally, while we have focused on LOCA and BCSD, it is important to note that they are both direct statistical downscaling methods based on GCM precipitation and temperature to project precipitation and temperature respectively; the differences between them are likely to be smaller than differences with other downscaling methods. If a GCM does not represent important aspects of the regional climate, the downscaled projected changes may not be reliable either; users are encouraged to consider the physical processes behind any hydro-climatic changes projected to understand and contextualize the reliability of those projections.

5. References

Alder, J. R., & Hostetler, S. W. 2019. The Dependence of Hydroclimate Projections in Snow-Dominated Regions of the Western United States on the Choice of Statistically Downscaled Climate Data. *Water Resources Research*, 55(3), 2279-2300. doi:10.1029/2018wr023458

Bohn, T.J., Livneh, B., Oyler, J.W., Running, S.W., Nijssen, B. and Lettenmaier, D.P., 2013. Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agricultural and forest meteorology*, *176*, pp.38-49, doi:10.1016/j.agrformet.2013.03.003.

Cannon, A.J., Sobie, S.R. and Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes?. *Journal of Climate*, *28*(17), pp.6938-6959, doi:10.1175/JCLI-D-14-00754.1

Fowler, H. J., S. Blenkinsop, and Tebaldi C., 2007. Review: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modeling, Int. J. Climatol., 27, 147–178, doi:10.1002/joc.1556.

Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T., Su, F., Sheffield, J., Pan, M., Lettenmaier, D. and Wood, E.F., 2010. Water Budget Record from Variable Infiltration Capacity (VIC) Model Algorithm Theoretical Basis Document.

Gutmann, E., Pruitt, T., Clark, M., Brekke, L., Arnold, J.R., Raff, D.A., and Rasmussen, R.M., 2014. An intercomparison of statistical downscaling methods used for water resource assessments in the United States, *Water Resources Research*, **50**, 7167-7186, doi:10.1002/2014WR015559.

Haerter, J., Hagemann, S., Moseley, C. and Piani, C., 2011. Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences*, *15*, pp.1065-1073, doi:10.5194/hess-15-1065-2011

Harding, B.L., A.W. Wood, and J.R. Prairie, 2012. The implications of climate change scenario selection for future streamflow projection in the upper Colorado River basin. *Hydrol. Earth Syst. Sci. Discuss.,* **9,** 847–894, doi:10.5194/hessd-9-847-2012.

Jiang, Y., Kim, J. B., Still, C. J., Kerns, B. K., Kline, J. D., & Cunningham, P. G. 2018. Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. *Sci Data*, 5, 180016. doi:10.1038/sdata.2018.16

Li, H. J. Sheffield, and E. F. Wood, 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile mapping. *J.Geophys. Res. D.*, v 115, No, D10101, doi:10.1029/2009JD012882, 2010.

Liang, X., Wood, E.F. and Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, *13*(1-4), pp.195-206, doi: 10.1016/0921-8181(95)00046-1.

Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E.P. and Lettenmaier, D.P., 2013. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *Journal of Climate*, *27*(1), pp.477-486, doi:10.1175/JCLI-D-12-00508.1.

Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E.P. and Lettenmaier, D.P., 2014. CORRIGENDUM: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions. *Journal of Climate*, *27*(1), pp.477-486.

Livneh, B., Bohn, T.J., Pierce, D.W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D.R. and Brekke, L., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the US, and Southern Canada 1950–2013. *Scientific data*, *2*, p.150042, doi:10.1038/sdata.2015.42.

Lukas, J., Barsugli, J., Doesken, N., Rangwala, I. and Wolter, K., 2014. Climate change in Colorado: a synthesis to support water resources management and adaptation. *University of Colorado, Boulder, Colorado*, doi:10.13140/RG.2.2.36741.35043.

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

Maraun, D., 2016. Bias correcting climate change simulations – a critical review. Current Climate Change Reports, v. 2, p. 211-220, doi:10.1007/s40641-016-0050-x.

Maurer, E. P., et al. 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. Journal of Climate 15(22): 3237-3251, doi:10.1175/JCLI-D-12-00508.1.

Maurer, E.P., L. Brekke, T. Pruitt, and P.B. Duffy, 2007. Fine-resolution climate projections enhance regional climate change impact studies, Eos Trans. AGU, 88(47), 504, doi:10.1029/2007EO470006.

Maurer, E.P. and Pierce, D.W., 2014. Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. *Hydrology and Earth System Sciences*, *18*(3), pp.915-925, doi:10.5194/hess-18-915-2014

Maurer, E. P., H. G. Hidalgo, T. Das, M. D. Dettinger, and D. R. Cayan, 2010: The utility of daily large-scale climate data in the assessment of climte change impacts no daily streamflow in California. Hydrol. Earth Syst. Sci., 14, 1125-1138, doi:10.5194/hess-14-1125-2010.

Mearns, L. O**.,** M. Bukovsky, S. Pryor, and V. Magaña, 2014: Downscaling of Climate Information. Chapter 5 in Ohring, G. (Ed.) *Climate Change in North America*, Springer: New York, pp. 201-250.

Mendoza, P.A., M. Clark, N. Mizukami, A. Newman, M. Barlage, E. Gutmann, R. Rasmussen, B. Rajagopalan, L. Brekke, and J. Arnold, 2015. Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. *J. Hydrometeor*, **16**, 762–780, doi:10.1175/JHM-D-14-0104.1

Michelangeli, P.A., Vrac, M. and Loukos, H., 2009. Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters*, *36*(11), doi:10.1029/2009GL038401

Mizukami, N., Clark, M. P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R. J., Markstrom, S. L., Hay, L. E., Woods, R., Arnold, J. R., and Brekke, L. D. 2016. mizuRoute version 1: a river network routing tool for a continental domain water resources applications, Geosci. Model Dev., 9, 2223–2238, doi:10.5194/gmd-9-2223-2016.

Panofsky, H.A. and Brier, G.W., 1968. Some applications of statistics to meteorology, 224 pp. *Pa. State Univ., University Park, Pa*.

Pierce, D.W., Cayan, D.R. and Thrasher, B.L., 2014. Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), pp.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, v. 16, p. 2421-2442. doi:10.1175/JHM-D-14-0236.1

Pierce, D.W., J.F. Kalansky, and D.R. Cayan. 2018. Climate, Drought, and Sea Level Rise Scenarios for the Fourth California Climate Assessment. *California's Fourth Climate Change Assessment. Publication number: CCCA4- CEC-2018-006.* California Energy Commission: Sacramento, CA, https://www.energy.ca.gov/sites/default/files/2019-11/Projections_CCCA4-CEC-2018-006_ADA.pdf.

Reclamation, 2011, West-Wide Climate Risk Assessments: Bias-Corrected and Spatially Downscaled Surface Water Projections, Technical Memorandum No. 86-68210-2011-01, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado, 138 p. https://www.usbr.gov/climate/secure/docs/2011secure/west-wide-climate-risk-assessments.pdf

Reclamation, 2013. Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs. U.S. Department of the Interior, Bureau of Reclamation, 104 p., available at: http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/downscaled_climate.pdf

 Reclamation, 2014. Downscaled CMIP3 and CMIP5 Hydrology Projections - Release of Hydrology Projections, Comparison with Preceding Information, and Summary of User Needs. U.S. Department of the Interior, Bureau of Reclamation,104 p., available at: https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/BCSD5HydrologyMemo.pdf

Reclamation, 2016. Report by Bracken, C., Downscaled CMIP3 and CMIP5 Climate Projections – Addendum, September 2016, available at: https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/Downscaled_Climate_Projections_Addendum_Sept2016.pdf

Sankarasubramanian, A., Vogel, R.M. and Limbrunner, J.F., 2001. Climate elasticity of streamflow in the United States. *Water Resources Research*, *37*(6), pp.1771-1781, doi:10.1029/2000WR900330.

Switanek, M. B., Troch, P. A., Castro, C. L., Leuprecht, A., Chang, H.-I., Mukherjee, R., and Demaria, E. M. C., 2017. Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes, *Hydrol. Earth Syst. Sci.*, 21, 2649–2666, doi:10.5194/hess-21-2649-2017

Thornton, P.E., Running, S.W., 1999. An improved algorithm for estimat- ing incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agric. For. Meteorol. 93 (4), 211–228, doi:10.1016/S0168-1923(98)00126-9.

USGCRP, 2018. *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* [Reidmiller, D.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, 1515 pp. doi: 10.7930/NCA4.2018.

Viger, R. J. 2014. Preliminary spatial parameters for PRMS based on the Geospatial Fabric, NLCD2001 and SSURGO, US Geological Survey, doi:10.5066/F7WM1BF7.

Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P. and Mearns, L.O., 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. *Supporting material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA*, *27, https://www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf*.

Wobus, C., Gutmann, E., Jones, R., Rissing, M., Mizukami, N., Lorie, M., Mahoney, H., Wood, A. W., Mills, D., Martinich, J. 2017. Climate change impacts on flood risk and asset damages within mapped 100-year floodplains of the contiguous United States. *Natural Hazards and Earth System Sciences*, 17(12), 2199-2211. doi:10.5194/nhess-17-2199-2017

Wood A.W., L.R. Leung, V. Sridhar, and D.P. Lettenmaier, 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change* **62**: 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

Appendix A. Supplemental Figures

Additional figures of historical values for temperature, precipitation, ET, runoff, and SWE, magnitudes (a) and percent change of differences (b)
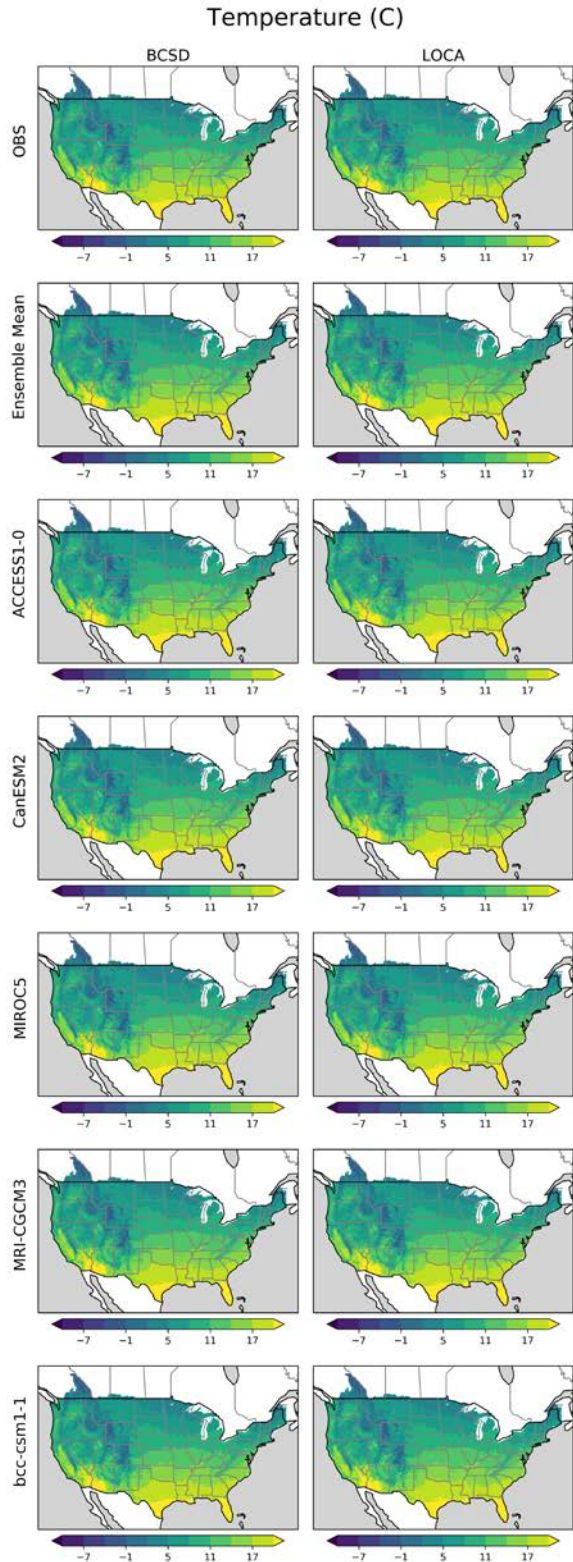
Figure A1a. Historical mean annual temperatures. Observations (top row) for both BCSD (left, Maurer) and LOCA (right, Livneh), the 23-model ensemble average (row 2), and five individual model examples (rows 3-7) averaged from 1970-1999. All 23 models (not shown) show little difference.
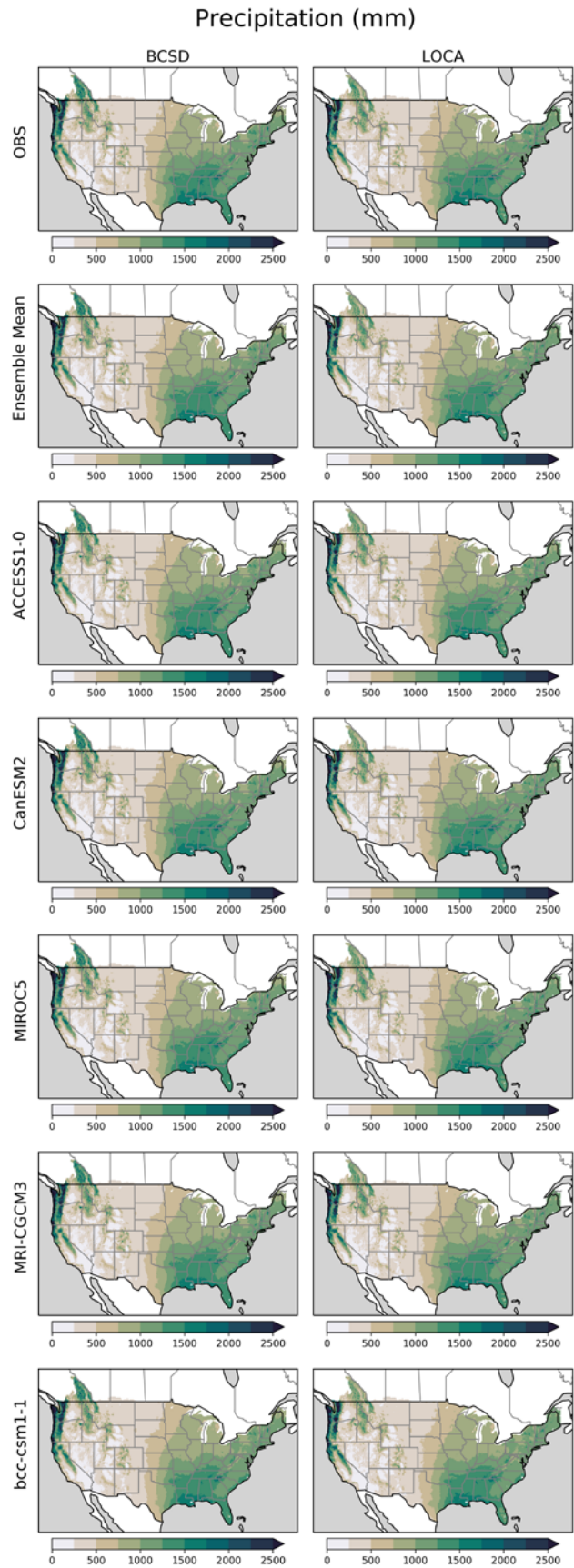
Precipitation (mm)

Figure A1b. Historical mean annual precipitation. Identical format to Figure A1a, except for precipitation instead of temperature. All 23 models (not shown) show little difference.
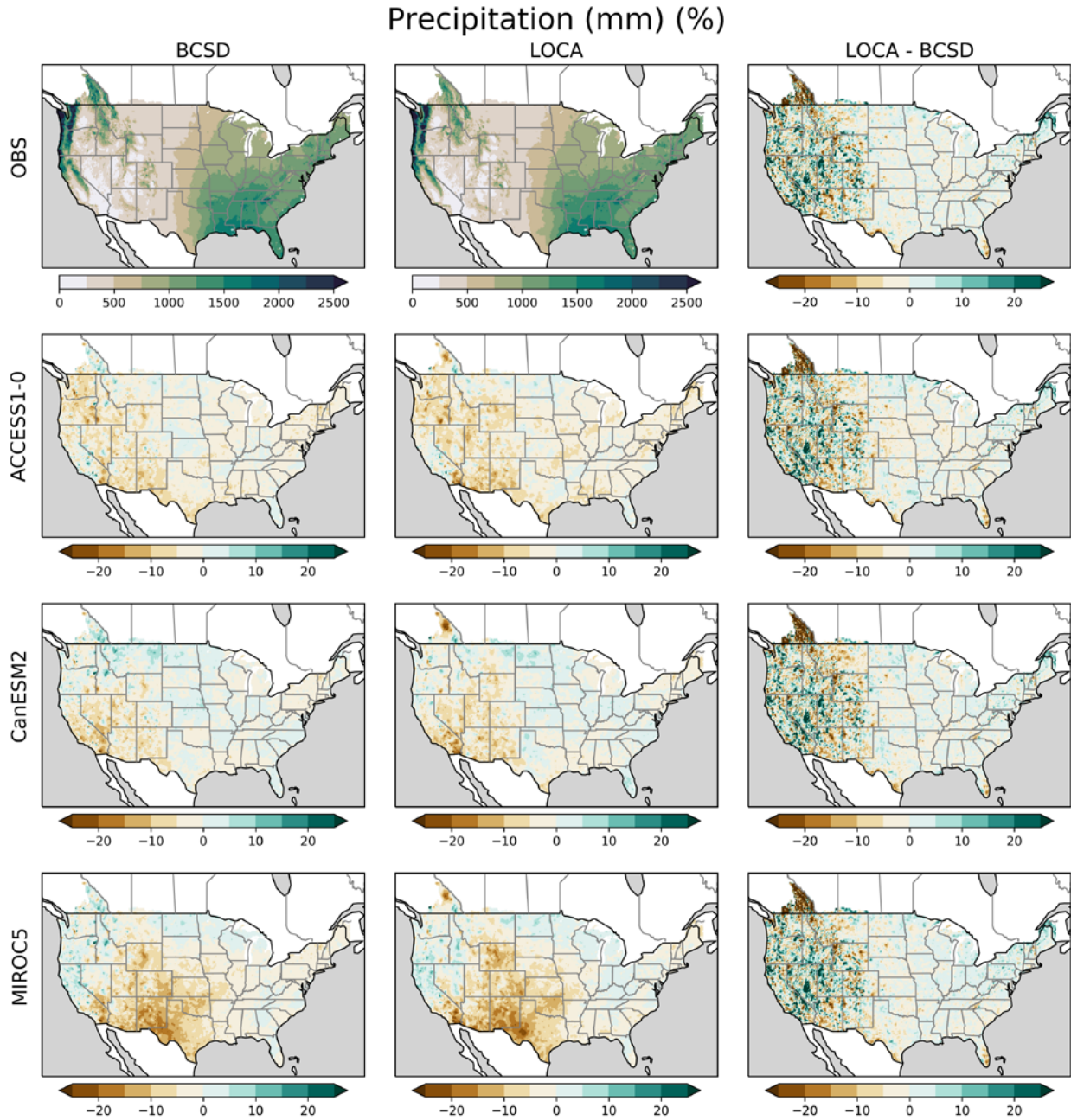
Figure A2b. Percent changes in precipitation. Identical to Figure 5, except the percentage change is plotted instead of total magnitude differences. The percent change is relative to OBS for the left (BSCD) and middle (LOCA) column of rows 2-4. The percent change of the difference (right column) is relative to BCSD (i.e. (LOCA-BCSD)/BCSD).
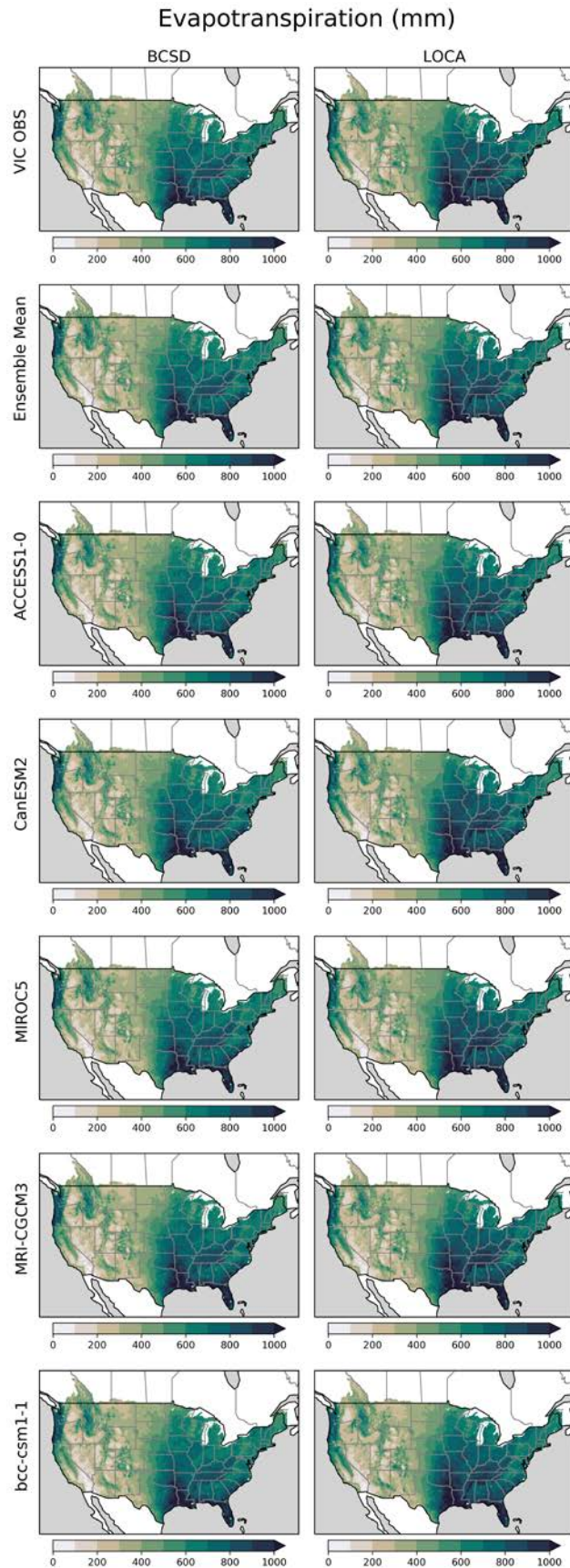
Figure A3a. Historical mean annual evapotranspiration. Values from VIC hydrologic model simulations run using observations (top row) from Maurer (BCSD) and Livneh (LOCA), the ensemble average of VIC simulations run using 23-model downscaled GCMs for BCSD (left) and LOCA (right) (row 2), and VIC simulations run using five individual downscaled GCM examples for BCSD (left) and LOCA (right) (rows 3-7). All values are averages of 1970-1999.

69
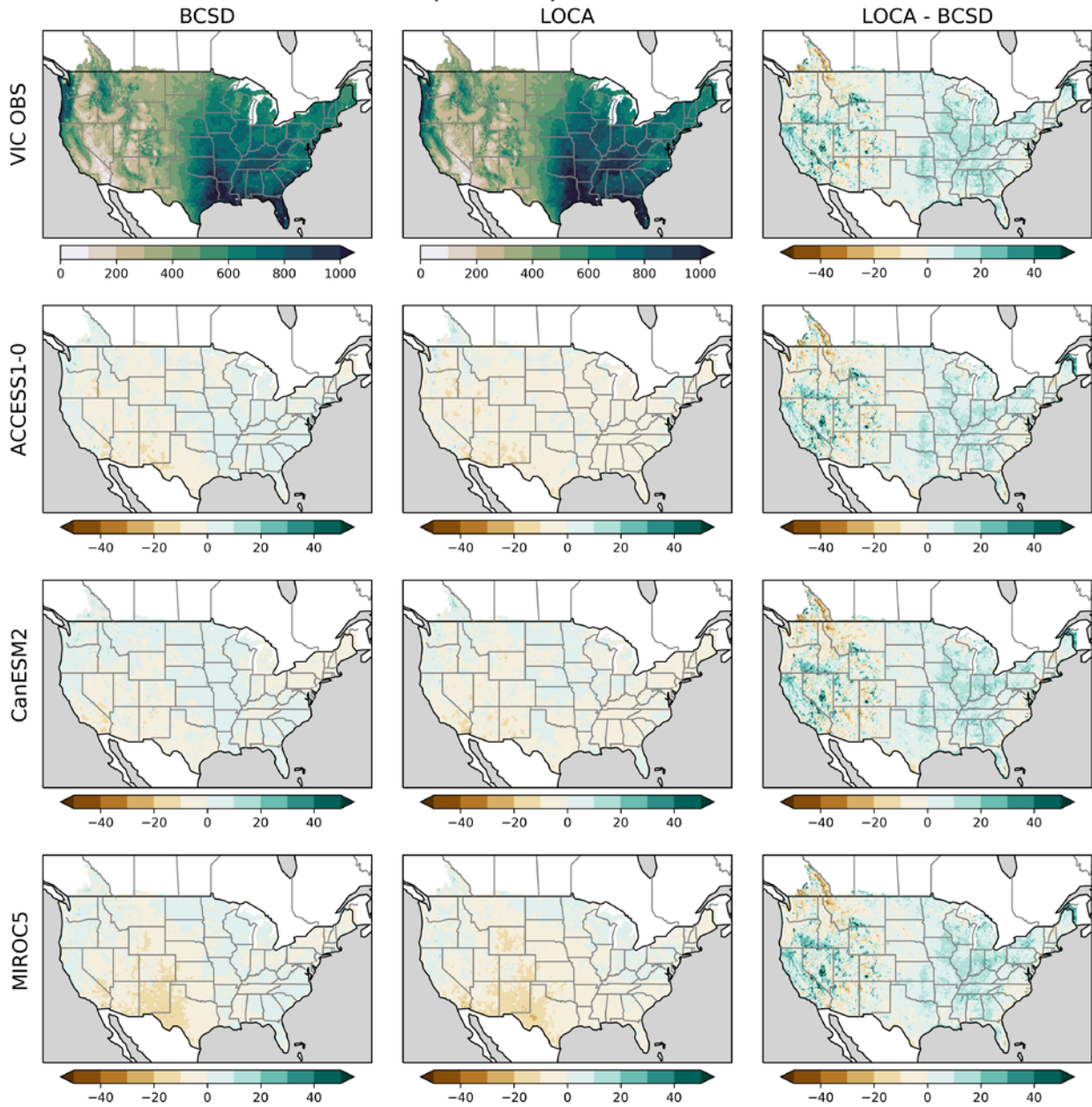
Figure A3b. Percent changes in ET. Identical to Figure 6, except the percentage change is plotted instead of total magnitude differences. The percent change is relative to VIC OBS for the left (BSCD) and middle (LOCA) column of rows 2-4. The percent change of the difference (right column) is relative to BCSD (i.e. (LOCA-BCSD)/BCSD).
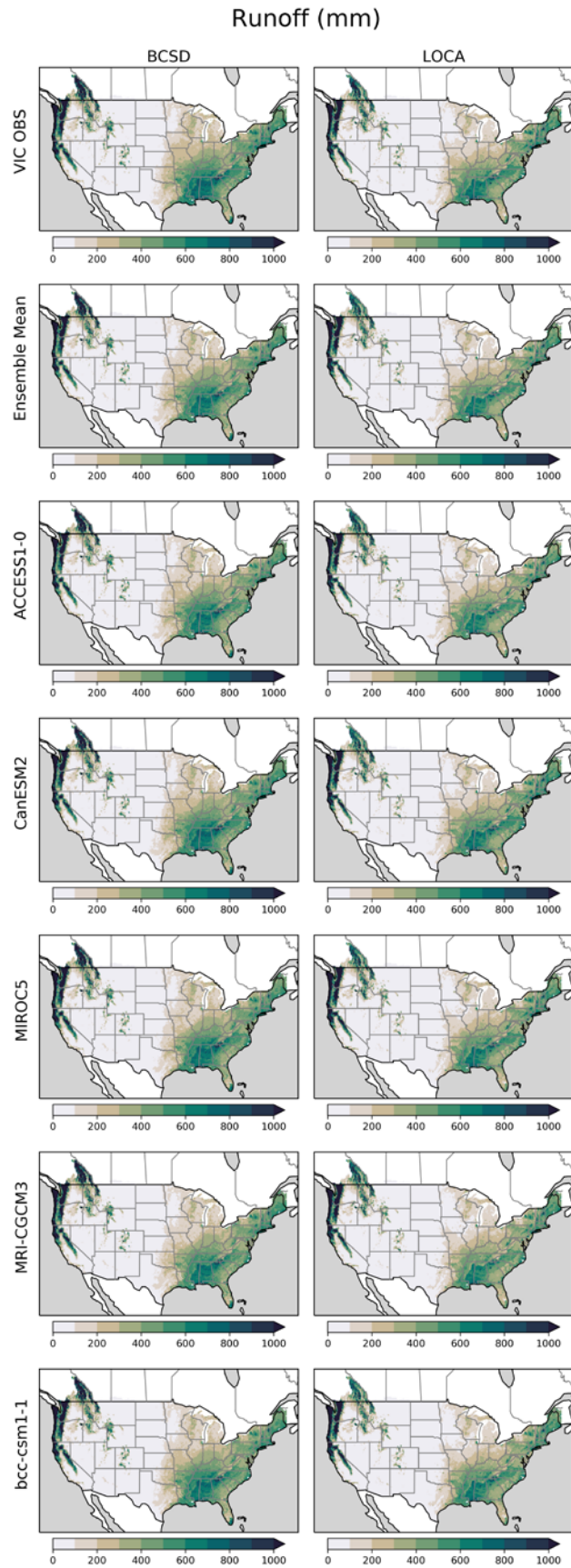
Figure A4a. Historical mean annual runoff.
Identical format to Figure A3a, except for VIC model output of runoff (surface runoff + baseflow) instead of ET.
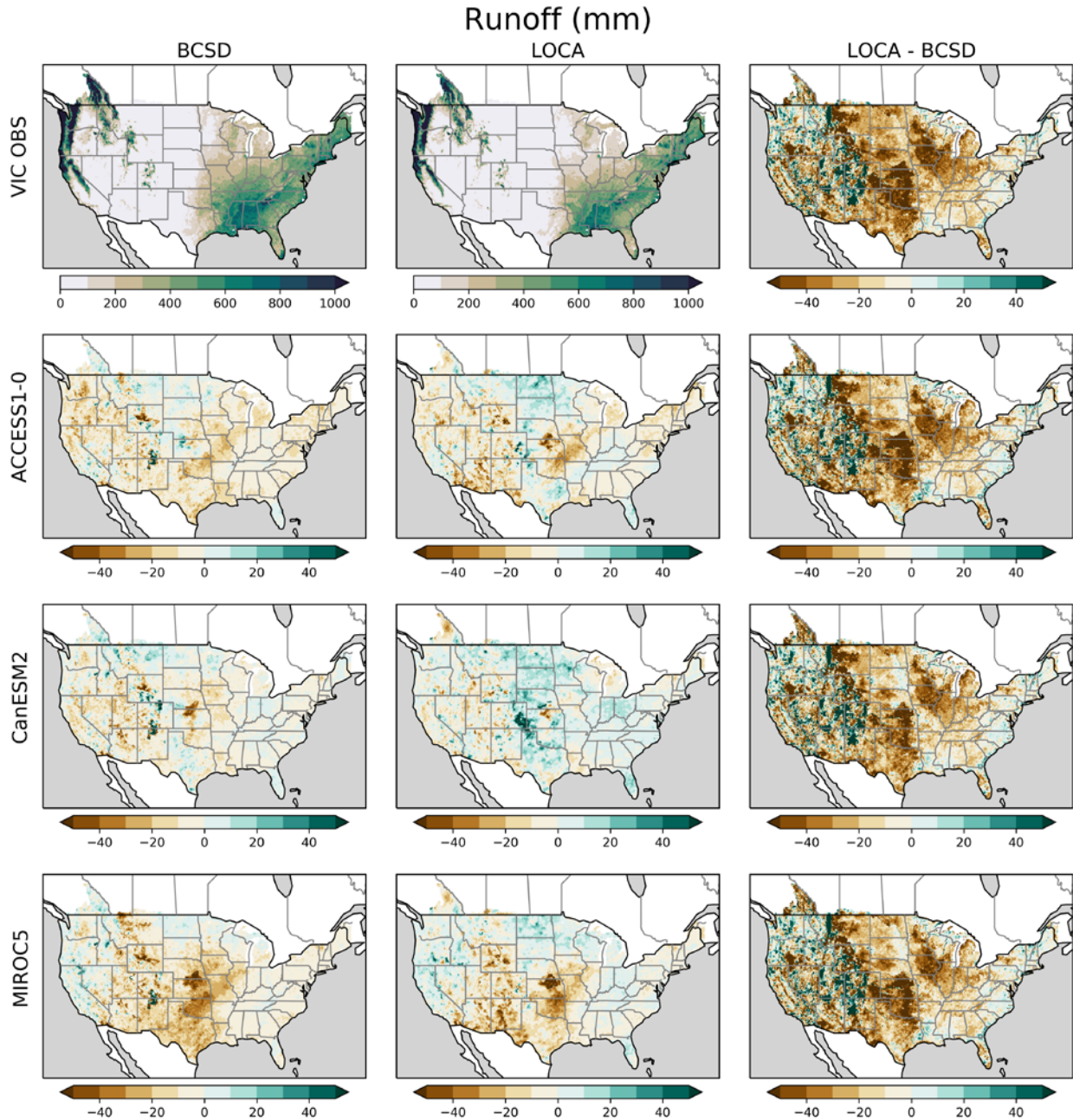
Figure A4b. Percent changes in runoff. Identical to Figure 7, except the percentage change is plotted instead of total magnitude differences. The percent change is relative to VIC OBS for the left (BSCD) and middle (LOCA) column of rows 2-4. The percent change of the difference (right column) is relative to BCSD (i.e. (LOCA-BCSD)/BCSD).

*Note percentages in locations with little runoff are likely to be considerably large. This is why we include this only in the appendix for those interested in percent changes in particular areas. Total magnitude changes are shown in Figure 7.

Snow Water Equivalent (mm)
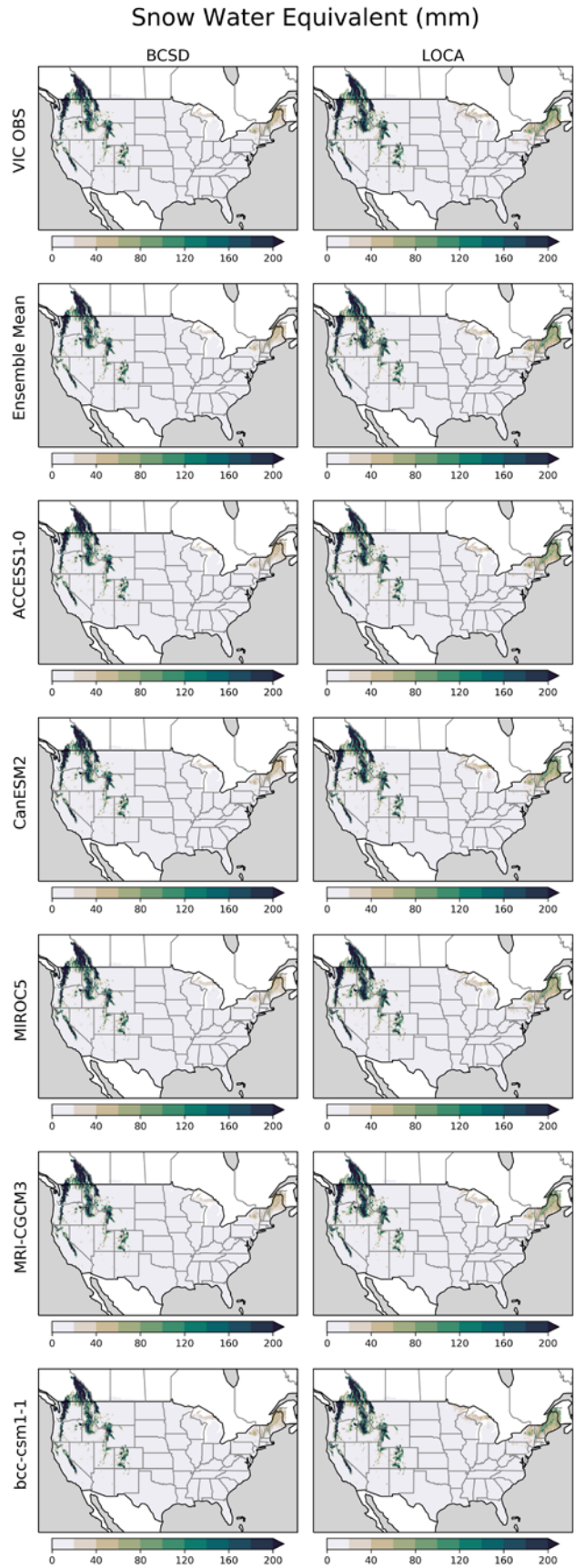
Figure A5a. Historical mean annual snow water equivalent (SWE). Identical format to Figure A3a, except for VIC model output of SWE instead of ET.
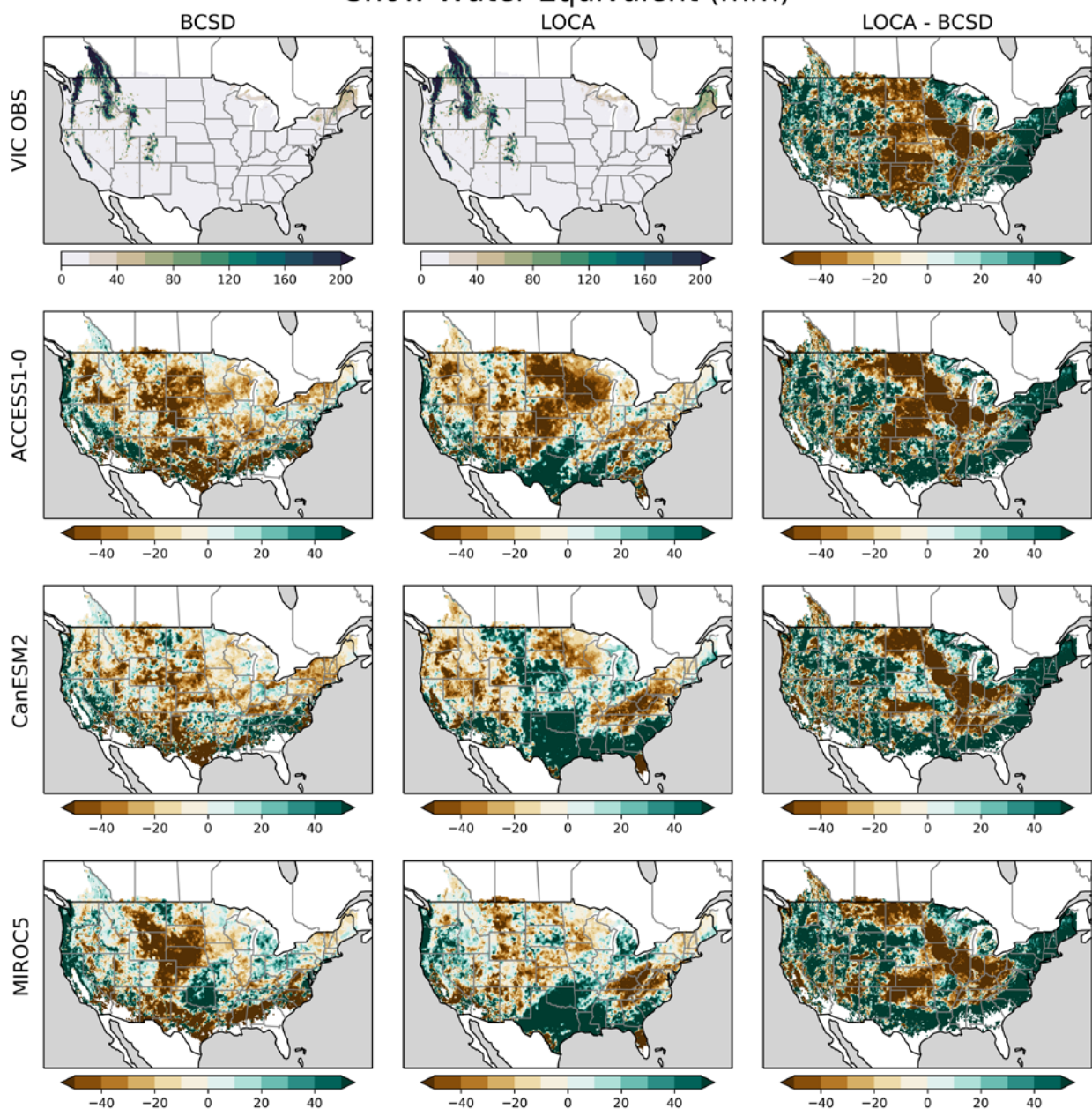
Figure A5b. Percent changes in snow water equivalent (SWE). Identical to Figure 8, except the percentage change is plotted instead of total magnitude differences. The percent change is relative to VIC OBS for the left (BSCD) and middle (LOCA) column of rows 2-4. The percent change of the difference (right column) is relative to BCSD (i.e. (LOCA-BCSD)/BCSD).

*Note percentages in locations with little SWE are likely to be considerably large. This is why we include this only in the appendix for those interested in percent changes in particular areas. Total magnitude changes are shown in Figure 8.

Appendix B. Interannual variability in precipitation

In addition to long-term means, we compared interannual variability depicted by LOCA and BCSD precipitation values. LOCA and BCSD have similar interannual variability as averaged across all GCMs (Figure B1, row 2 Ensemble Mean), with LOCA having slightly more interannual variability across much of the CONUS. However, LOCA exhibits less spread across models in the interannual variability for different GCMs (Figure B1, row 3 Ensemble Std). This is likely due to the frequency dependent bias correction employed in LOCA which explicitly tries to correct time variance. In the historical period, LOCA appears to have more spread in interannual variability across models on average in the PNW (blue in the right column, 3 row sub-figure) and BCSD might have more in portions of California and the Canadian portion of the Columbia River basin (brown). These comparisons are noisy, however.

Notably, the differences between observations (top row) have a different pattern over much of the CONUS than the differences seen in individual GCMs, indicating that downscaled LOCA/BCSD differences are not determined exclusively by differences between the Livneh and Mauer training observation datasets.

Precipitation (mm), Interannual Variability
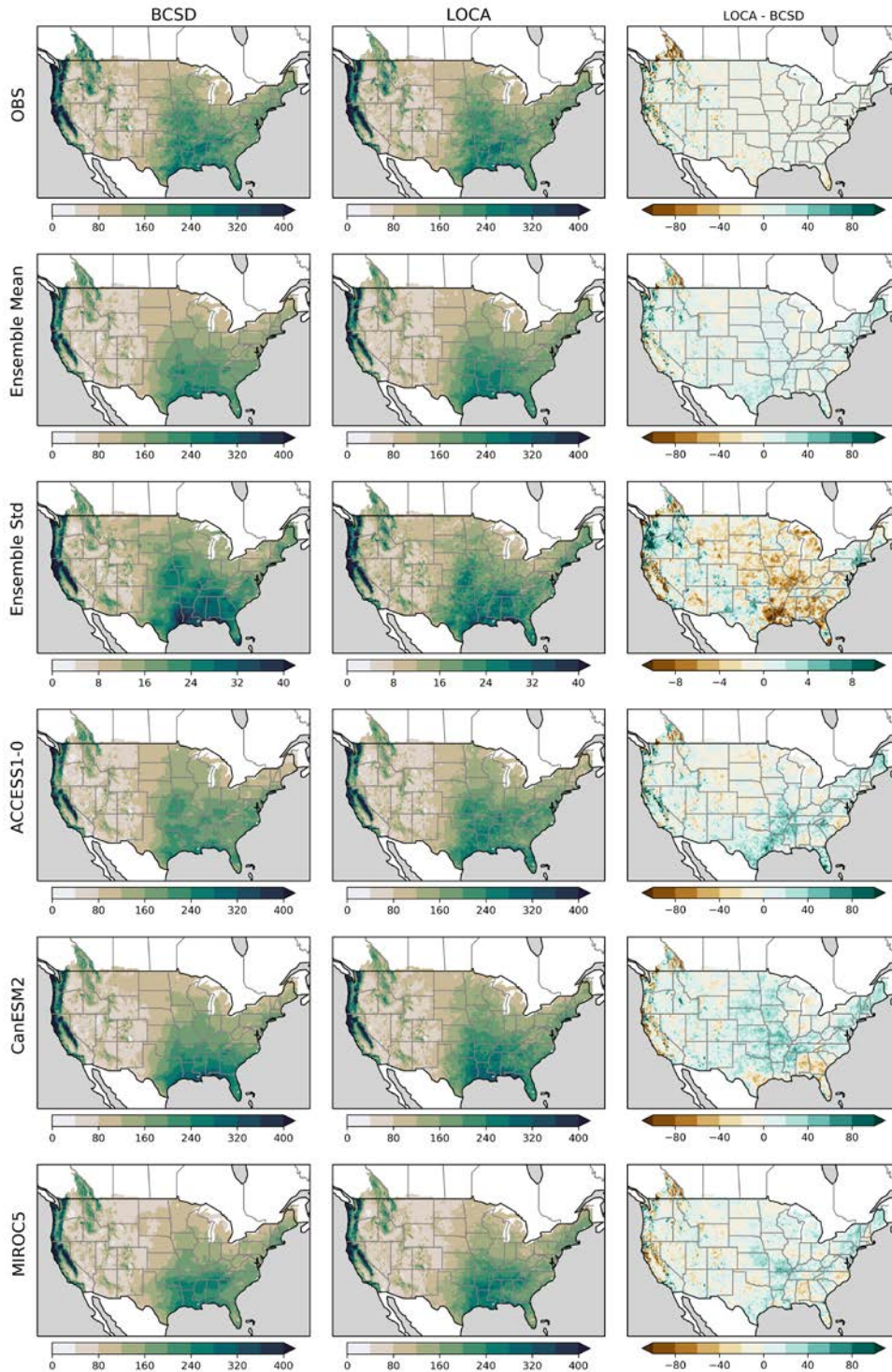
Figure B1. Interannual variability of precipitation. The standard deviation of annual precipitation from 1970-1999 for BCSD (left), LOCA (middle) and LOCA-BCSD (right) from observations (top row) averaged across all GCMs (row 2), and the standard deviation across GCMs (row 3), and three individual GCMS (rows 4-6).   Note: the color scale for standard deviation is a factor of 10 smaller.

Appendix C. Future changes for a subset of individual GCMs for 4.5 (a) and 8.5 (b)
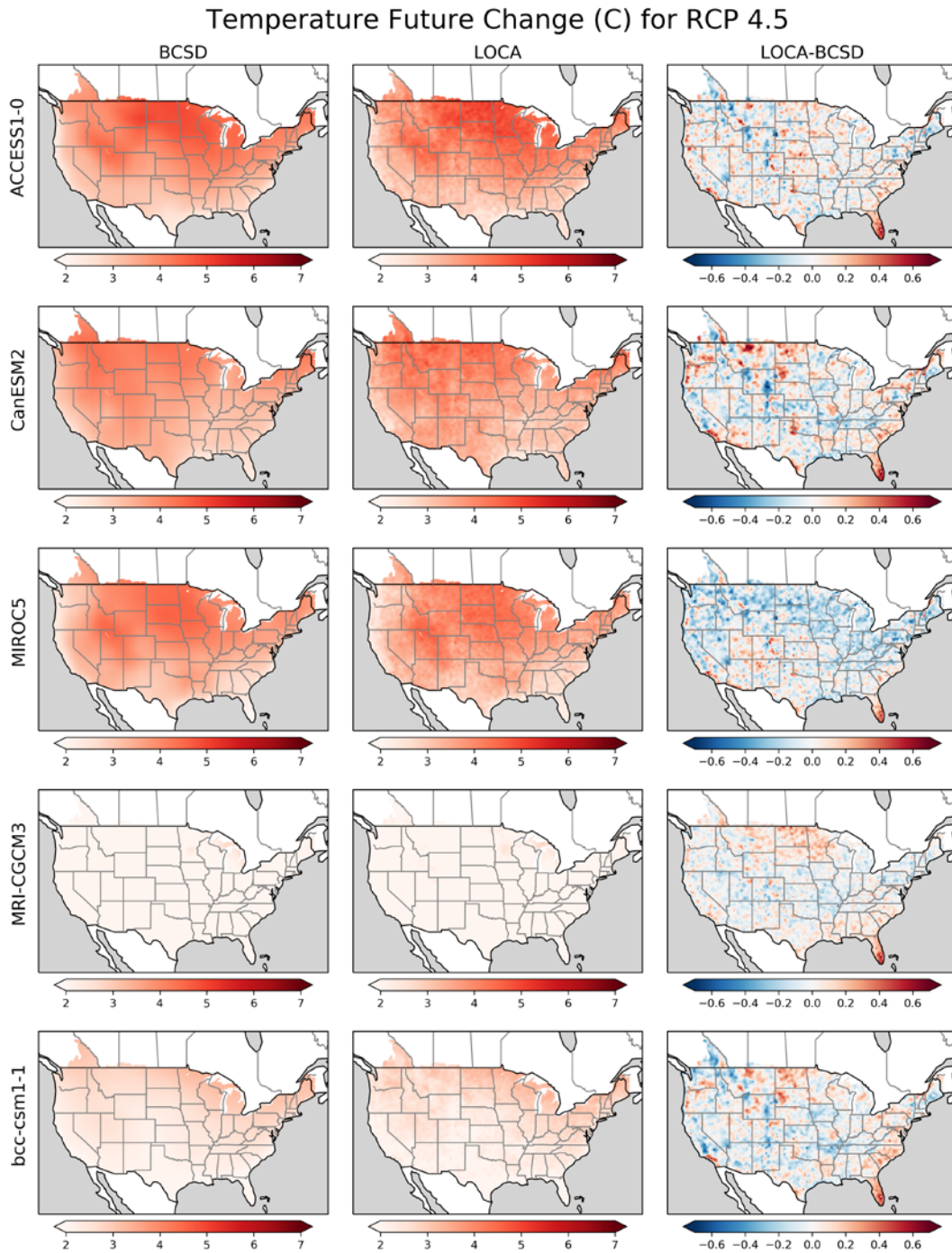
Figure C1a. Change in temperature for individual downscaled GCMs for RCP 4.5 Differences between the historical period (1970-1999) and future period (2070-2099) for RCP 4.5 for BCSD (left) and LOCA (middle) and the differences between LOCA and BCSD (right). Note colorbar is a factor of 10 smaller in the differences column (right).

Temperature Future Change (C) for RCP 8.5
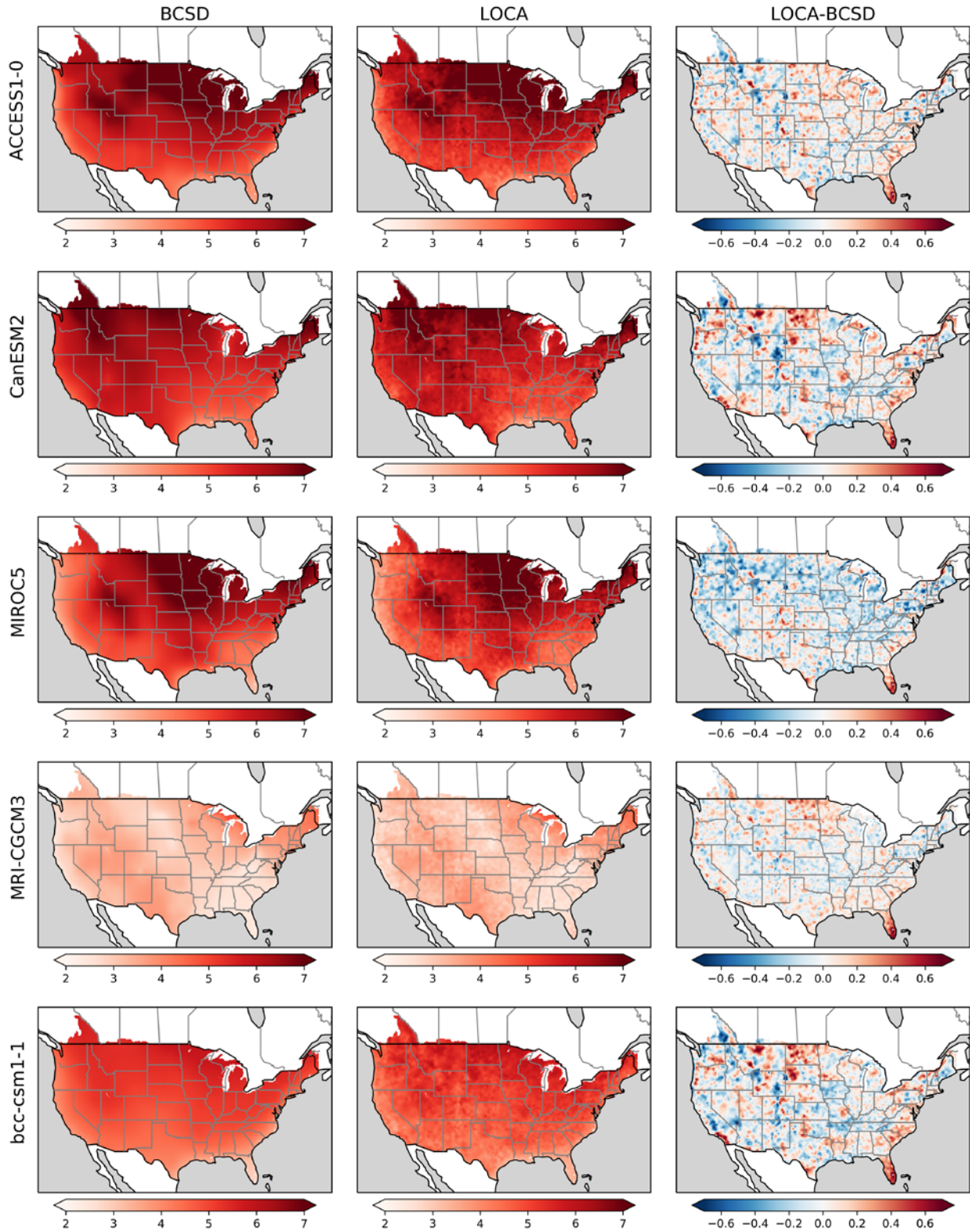
Figure C1b. Change signals in temperature for individual downscaled GCMs for RCP 8.5. Similar to previous figure but for 8.5 instead of 4.5.

Precipitation Future Change (mm) for RCP 4.5
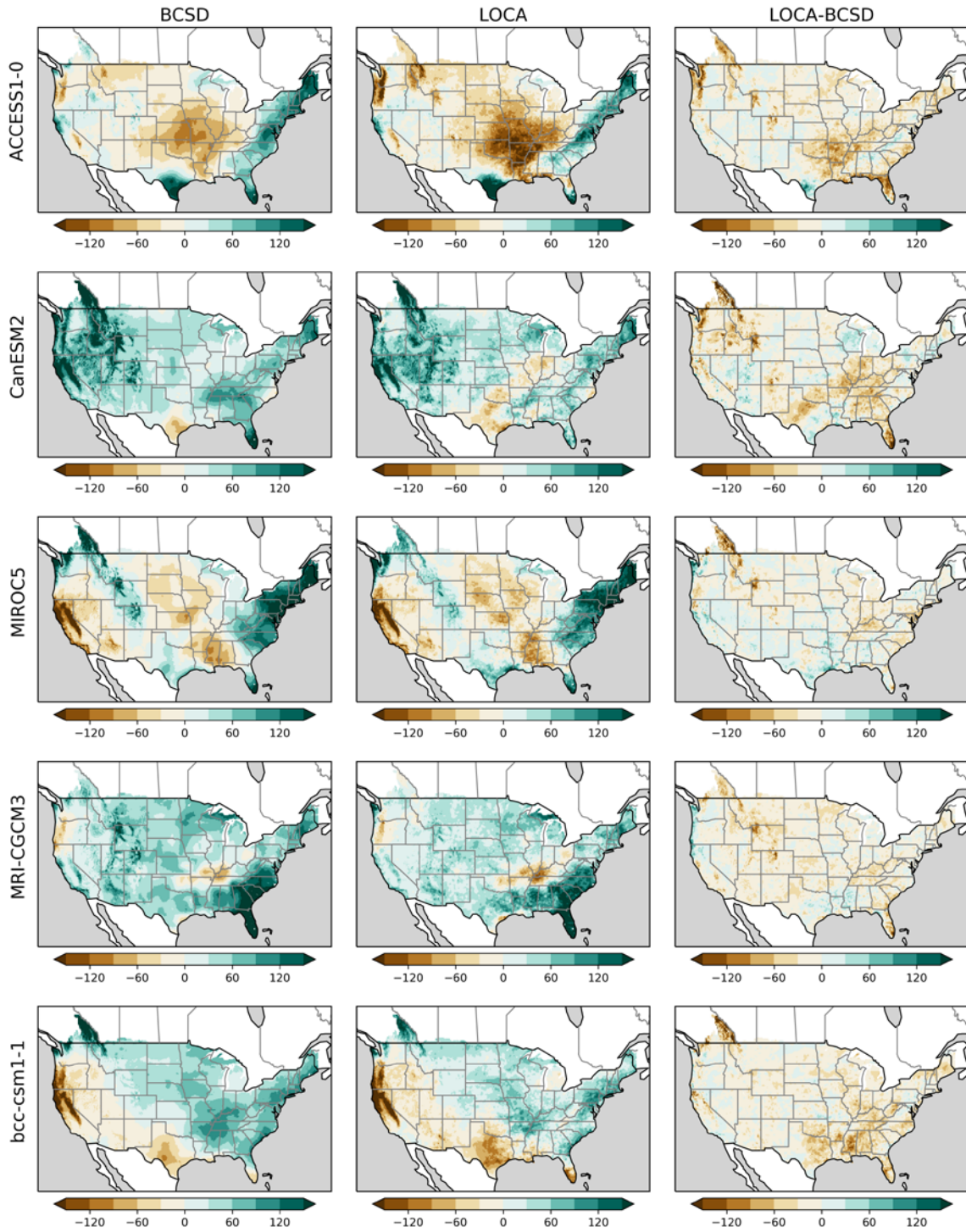
Figure C2a. Change signals in precipitation for individual downscaled GCMs for RCP 4.5.
Differences between the historical period (1970-1999) and future period (2070-2099) for RCP 4.5 for
BCSD (left) and LOCA (middle) and the differences between LOCA and BCSD (right).

Precipitation Future Change (mm) for RCP 8.5
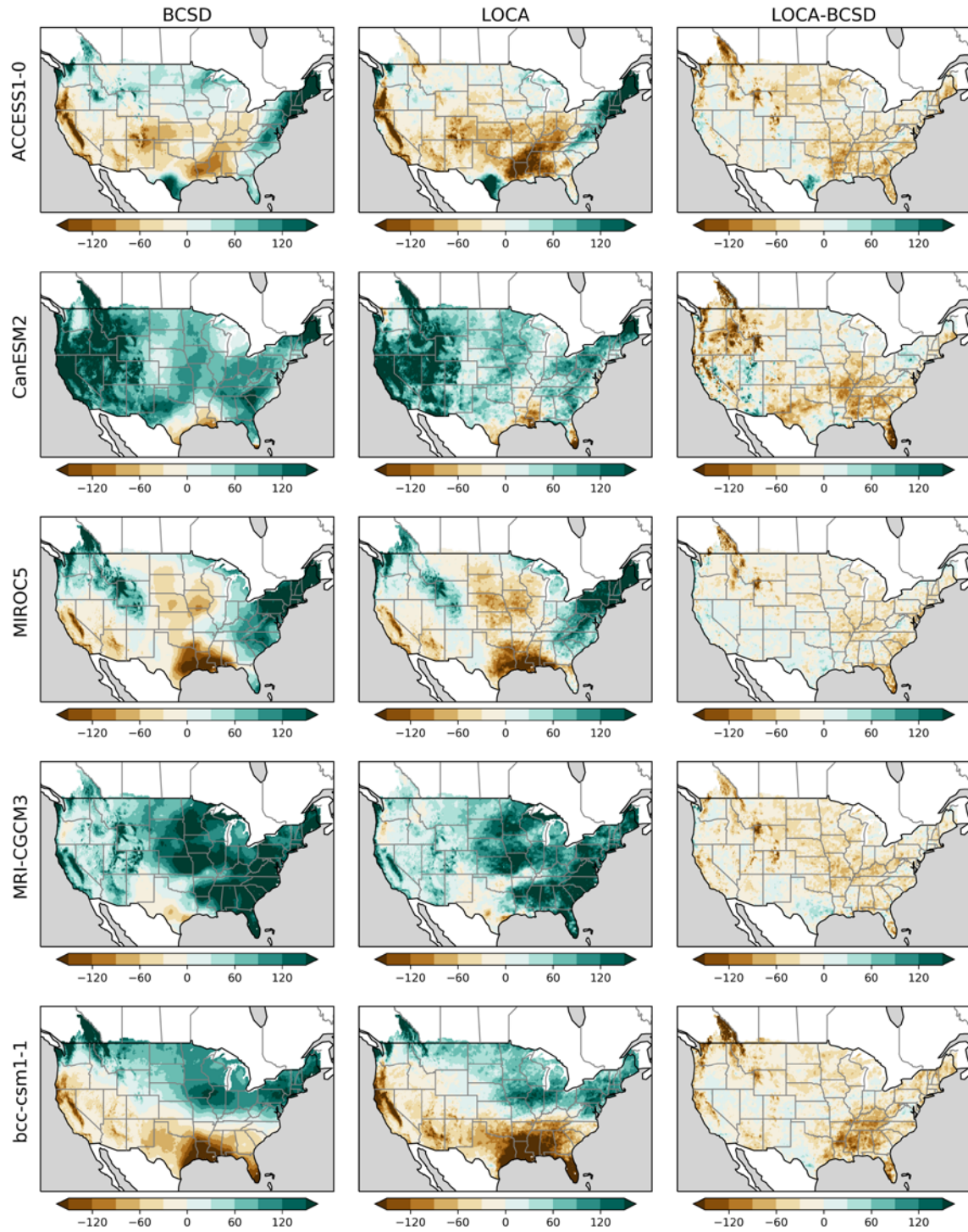
Figure C2b. Change signals in precipitation for individual downscaled GCMs for RCP 8.5. Similar to previous figure but for 8.5 instead of 4.5.
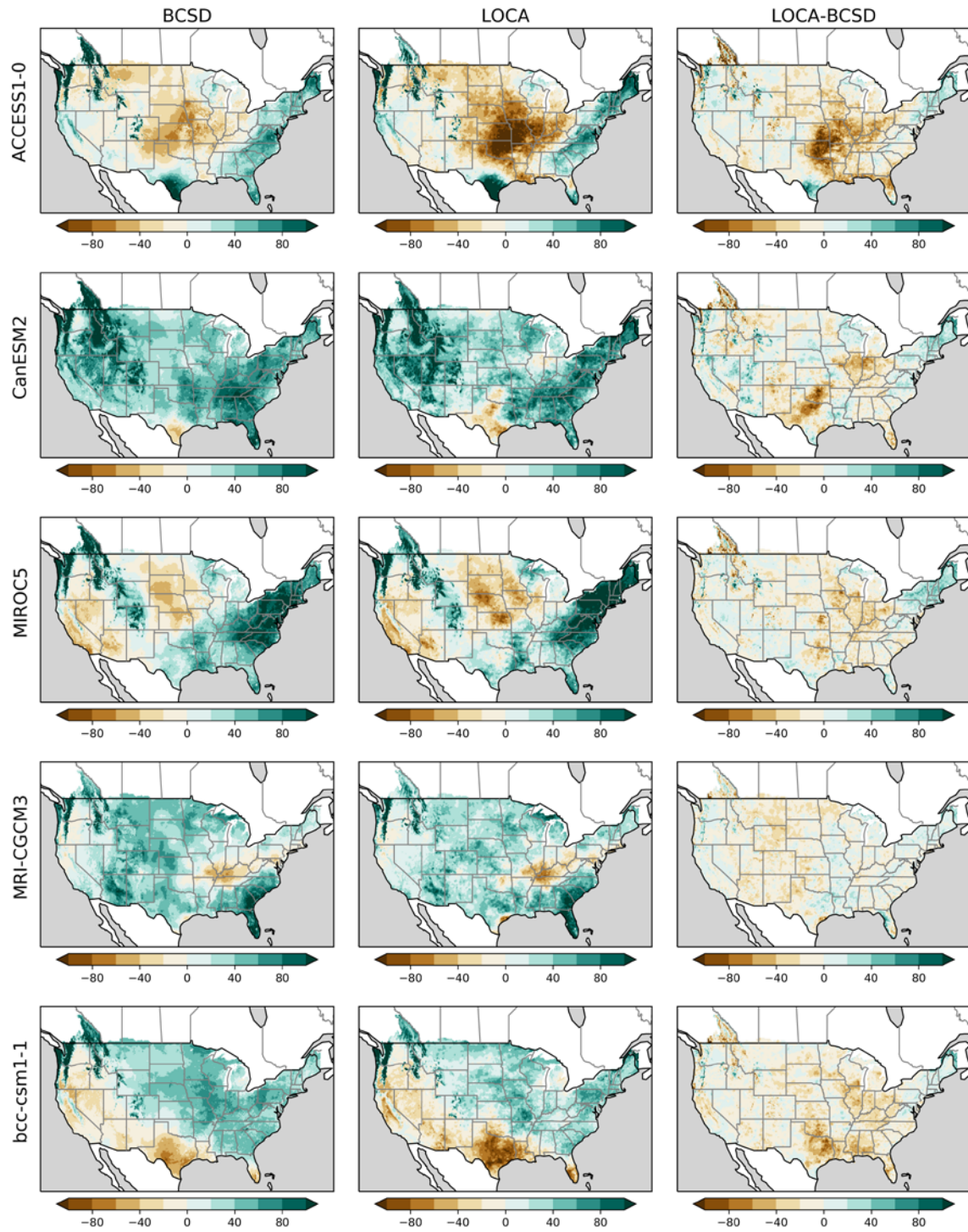
Figure C3a. Change in evapotranspiration for individual downscaled GCMs for RCP 4.5. Similar format to figure above but for ET instead of precipitation (note: color scale range is smaller here than it was for precipitation).
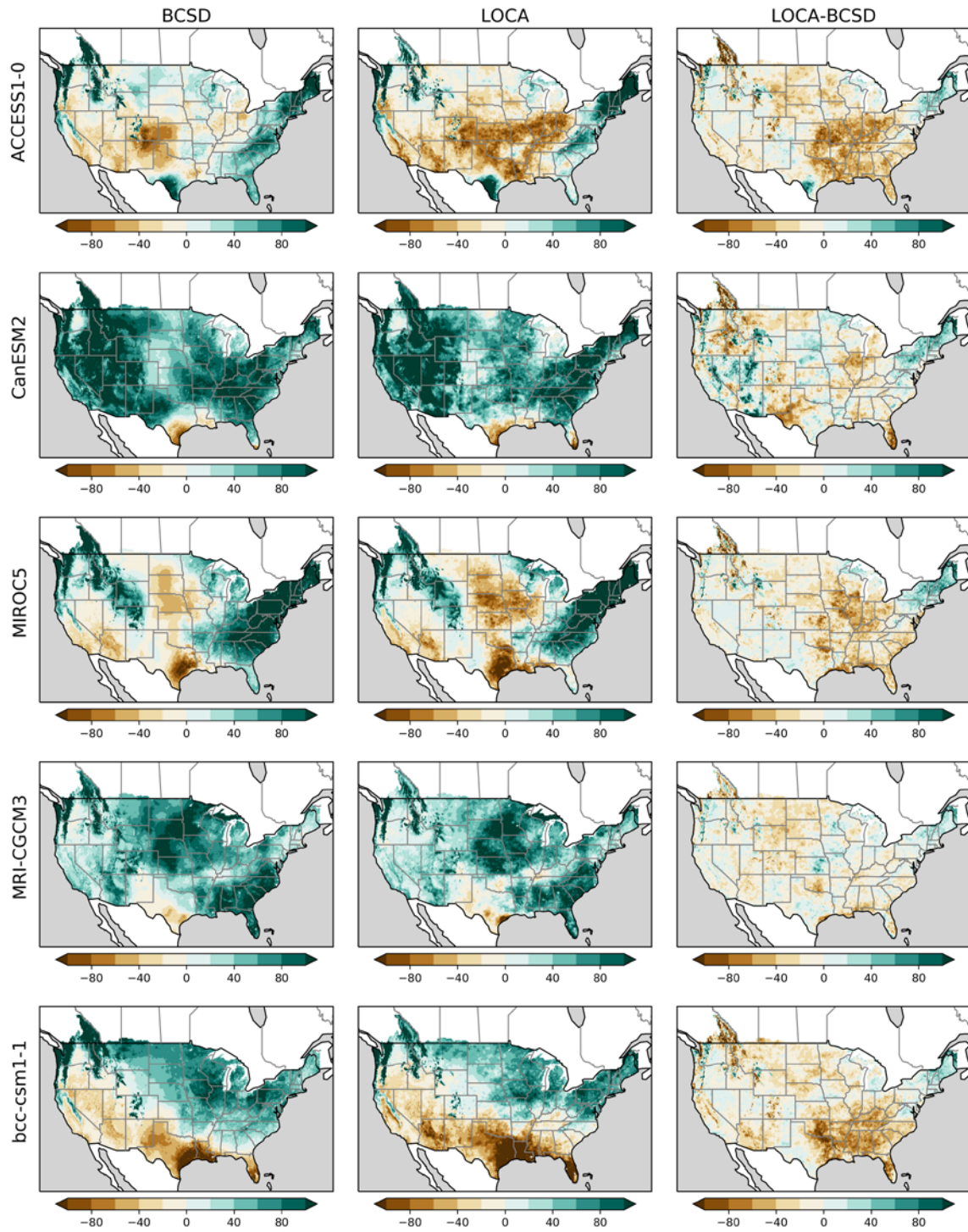
Figure C3b. Change signals in evapotranspiration for individual downscaled GCMs for RCP 8.5. Similar to previous figure but for 8.5 instead of 4.5.

Runoff Future Change (mm) for RCP 4.5
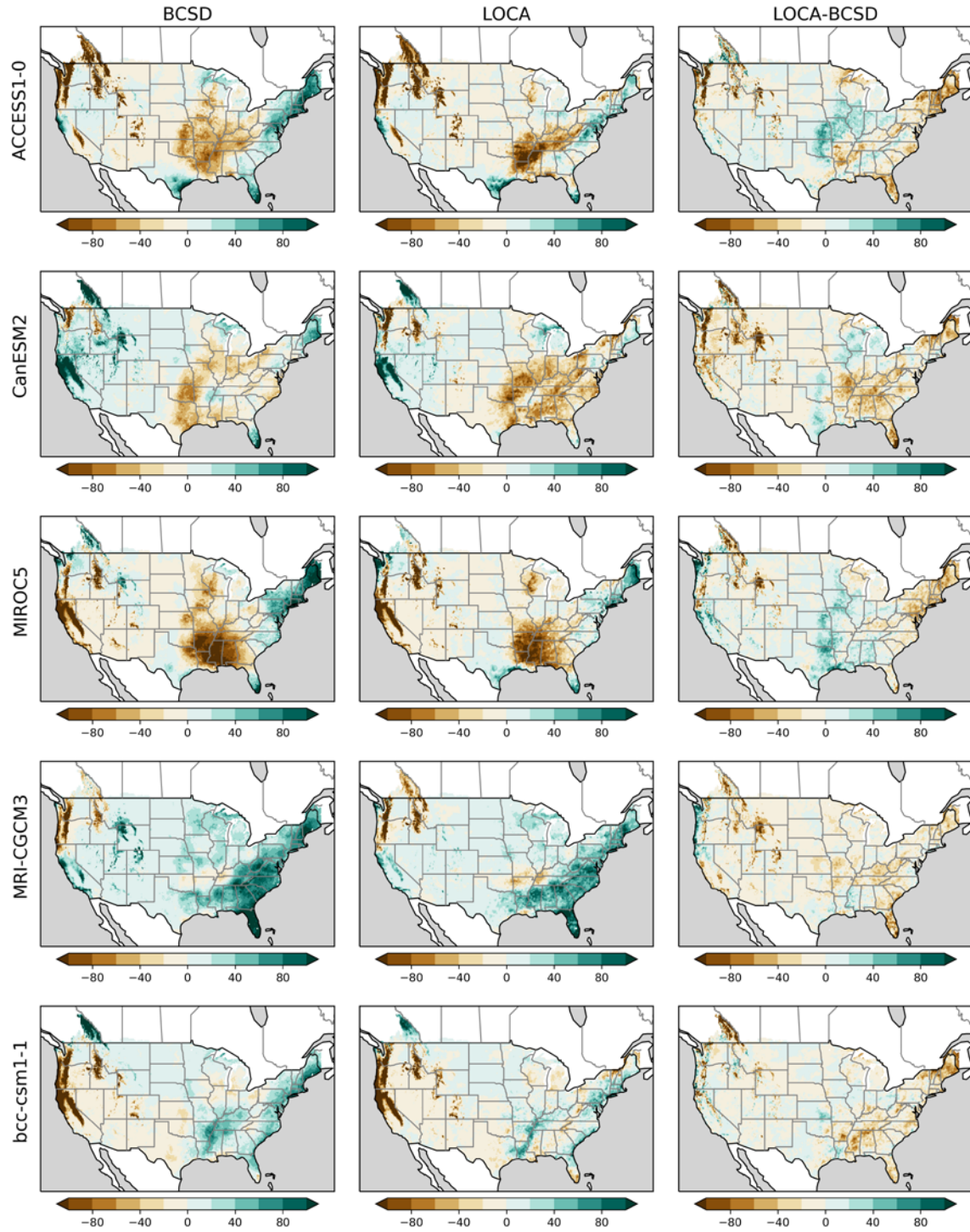
Figure C4a. Change in runoff for individual downscaled GCMs for RCP 4.5. Similar format to figure above but for runoff (surface runoff + baseflow) instead of ET.
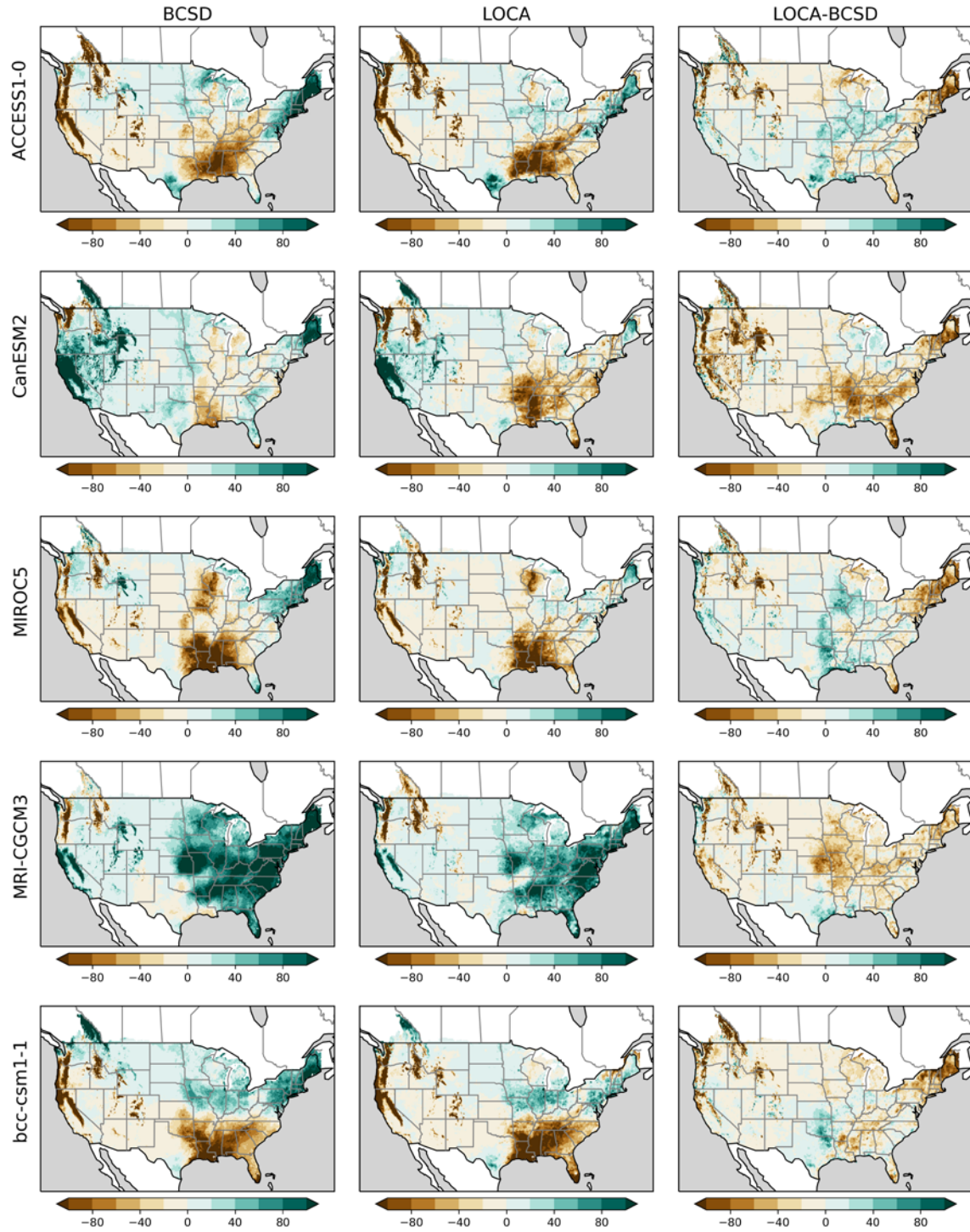
Figure C4b. Change signals in runoff for individual downscaled GCMs for RCP 8.5. Similar to previous figure but for 8.5 instead of 4.5.

SWE Future Change (mm) for RCP 4.5
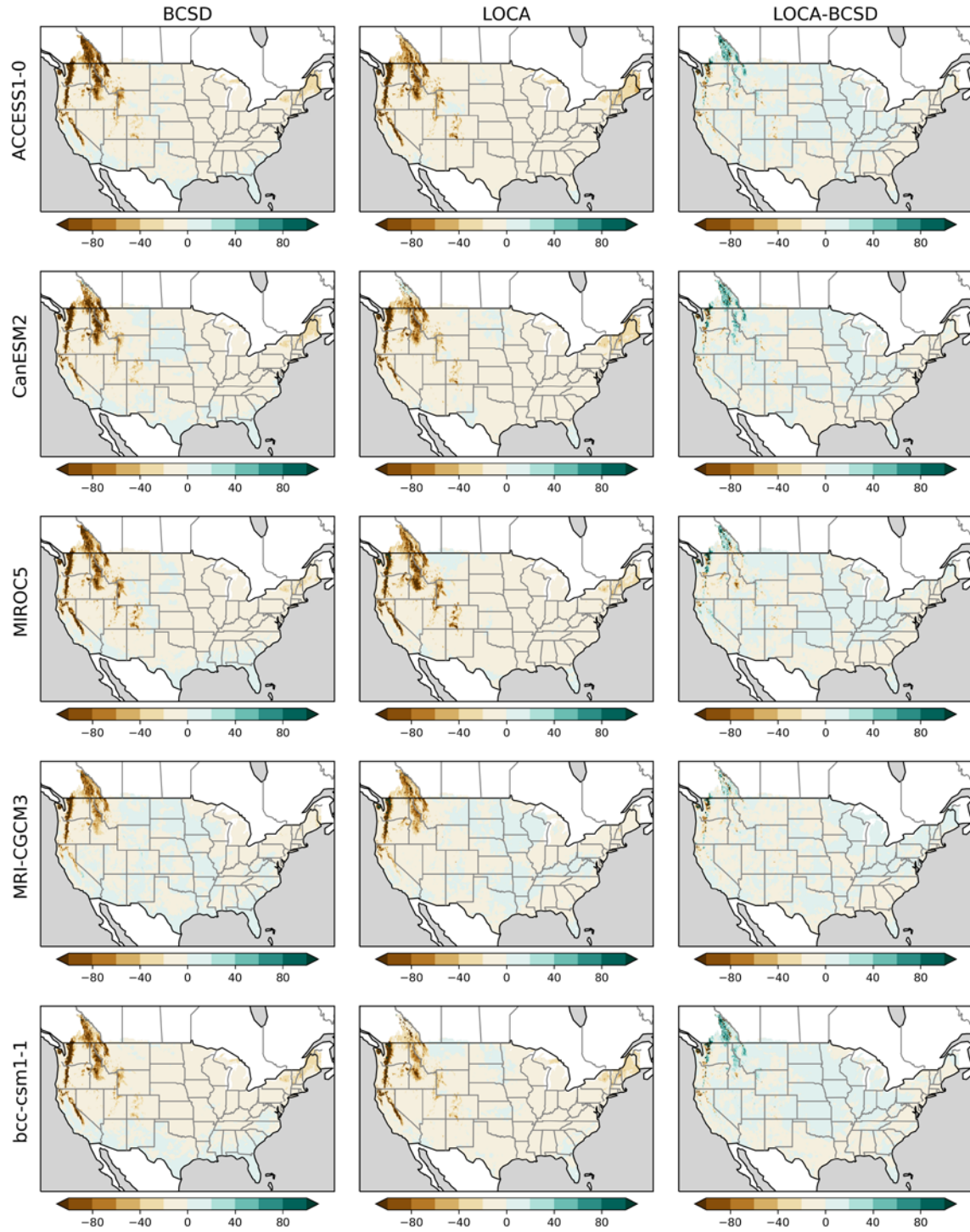
Figure C5a. Change signals in SWE for individual downscaled GCMs for RCP 4.5. Similar format to figures above but for SWE instead of runoff.
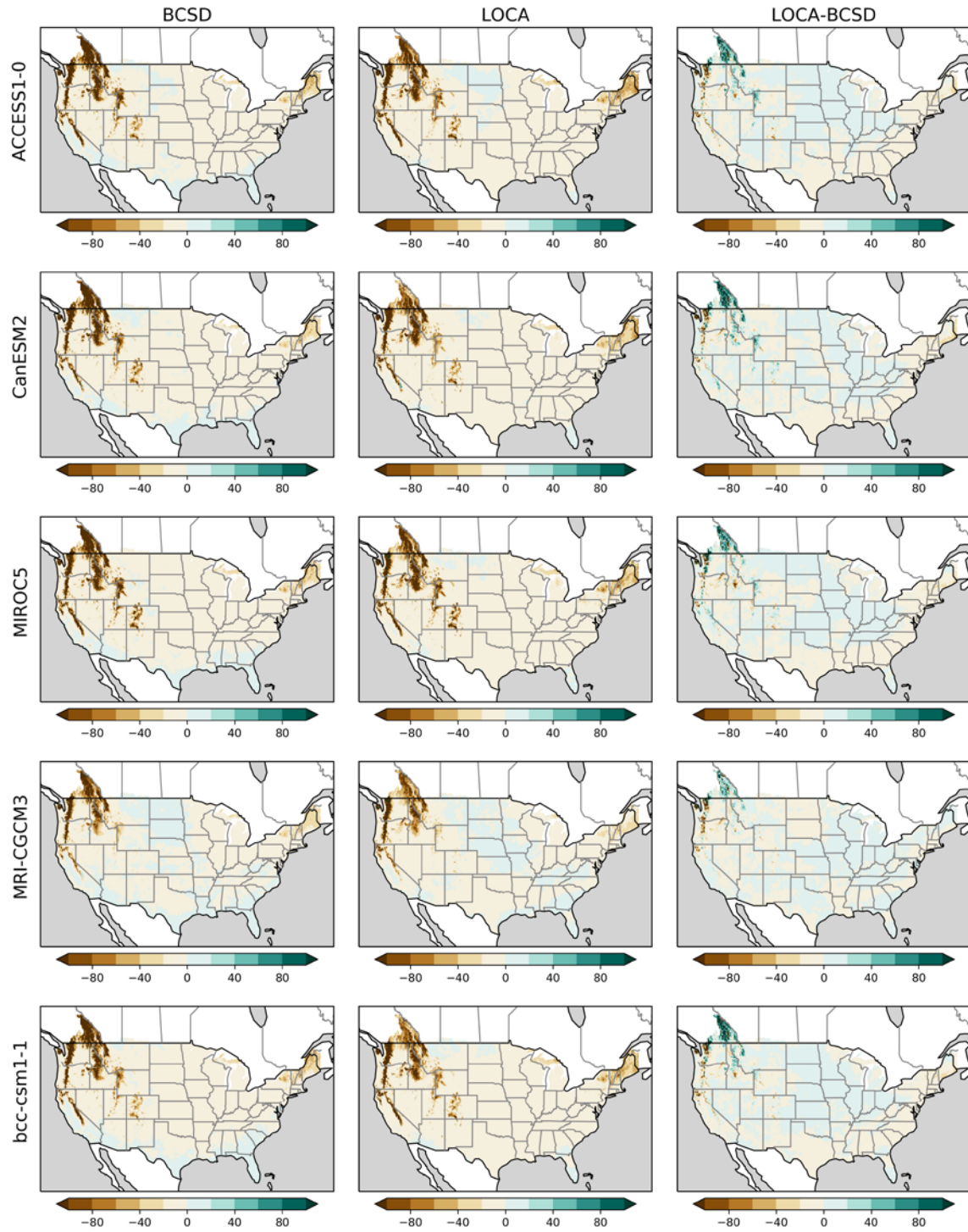
Figure C5b. Change in SWE for individual downscaled GCMs for RCP 8.5. Similar to previous figure but for 8.5 instead of 4.5.

Appendix D. Seasonal simulation of historical and future changes in all 18 HUC2
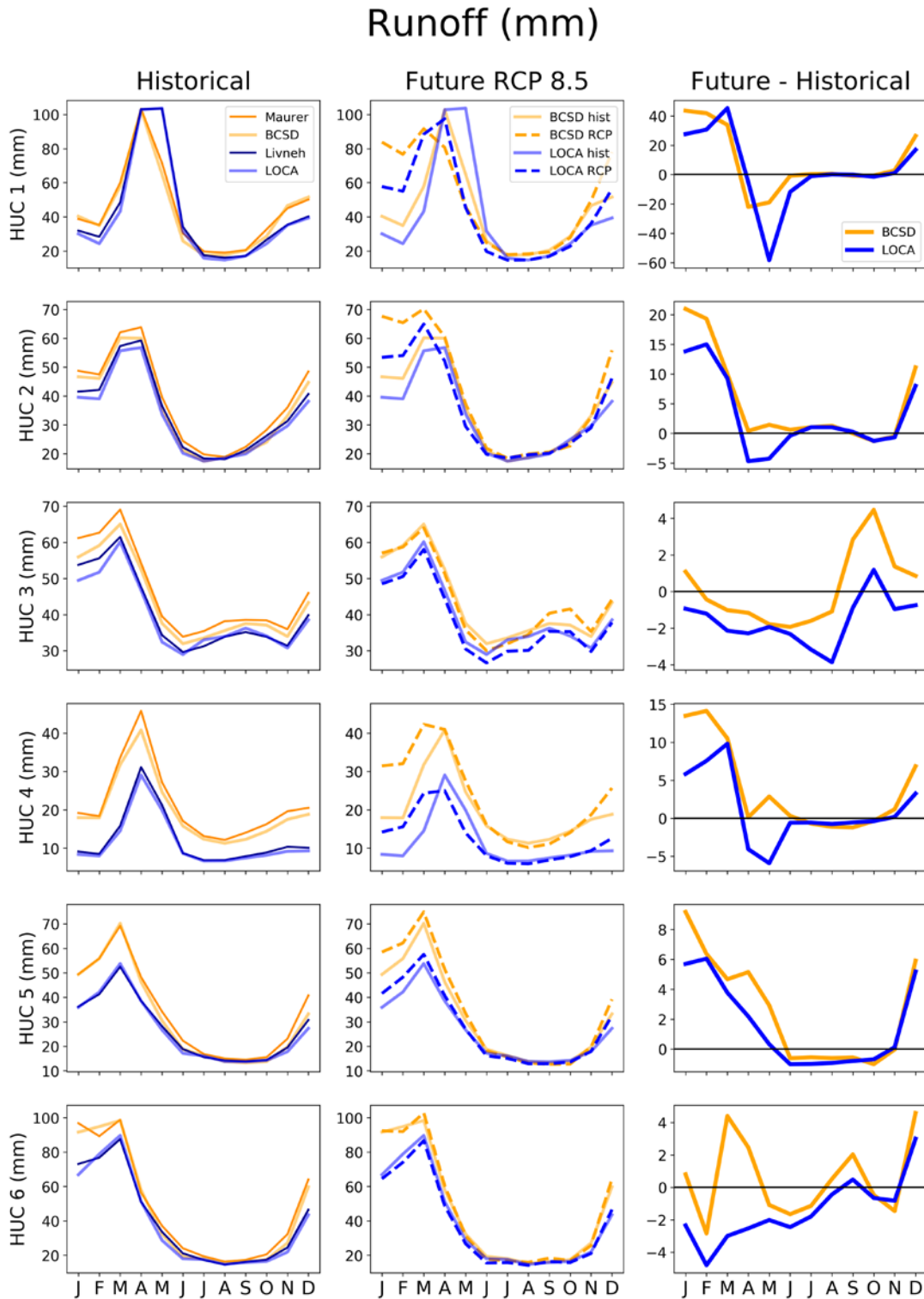
Runoff (mm)

Figure D1. Historical and future basin runoff for HUC2 01-06. Similar to Figure 16 but for but HUC2 01-06. Values are in mm, averaged across the watershed area.

Runoff (mm)
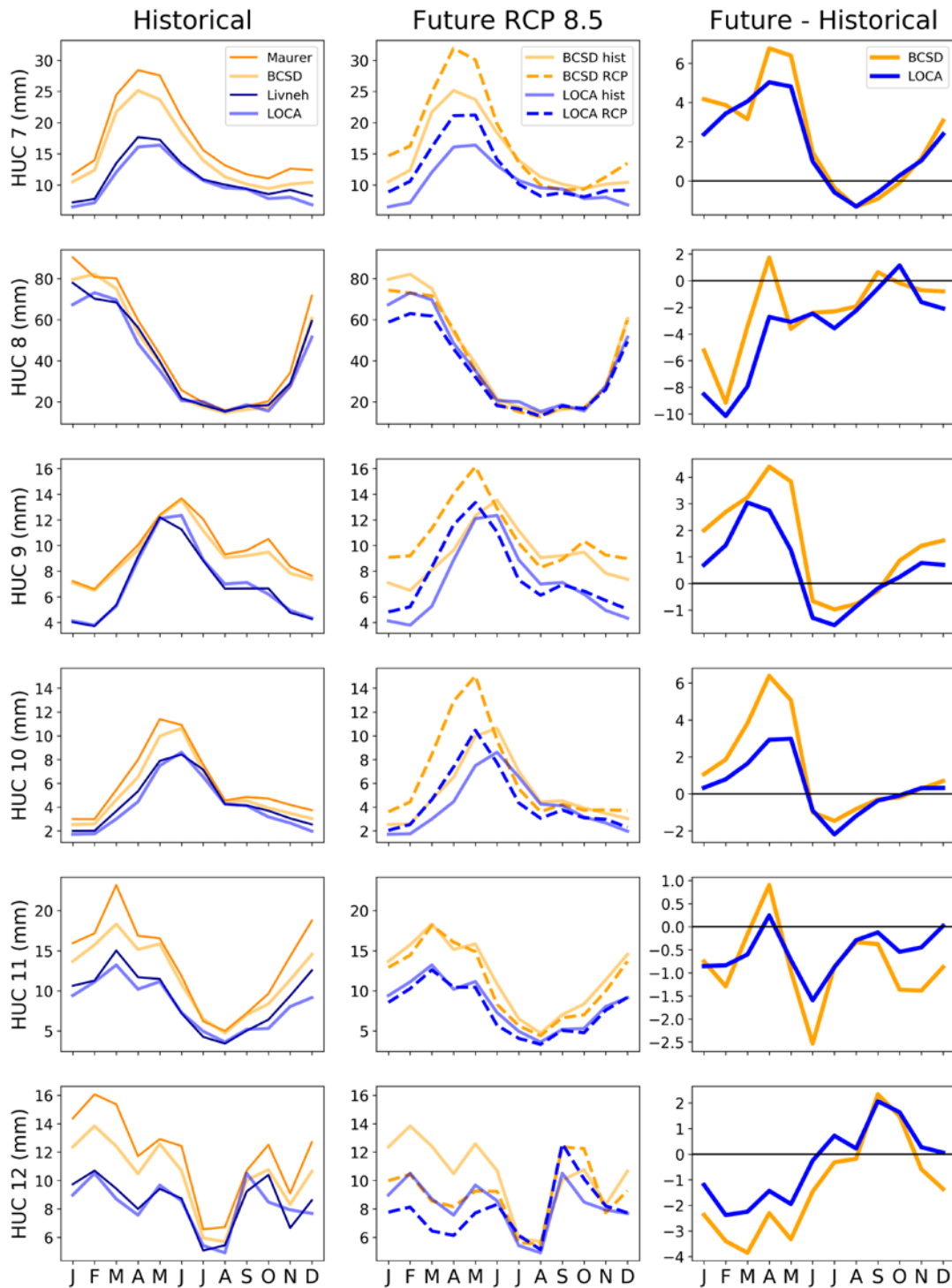
Figure D2. Historical and future basin runoff for HUC2 07-12. Similar to Figure 16 but for but HUC2 07-12. Values are in mm, averaged across the watershed area.
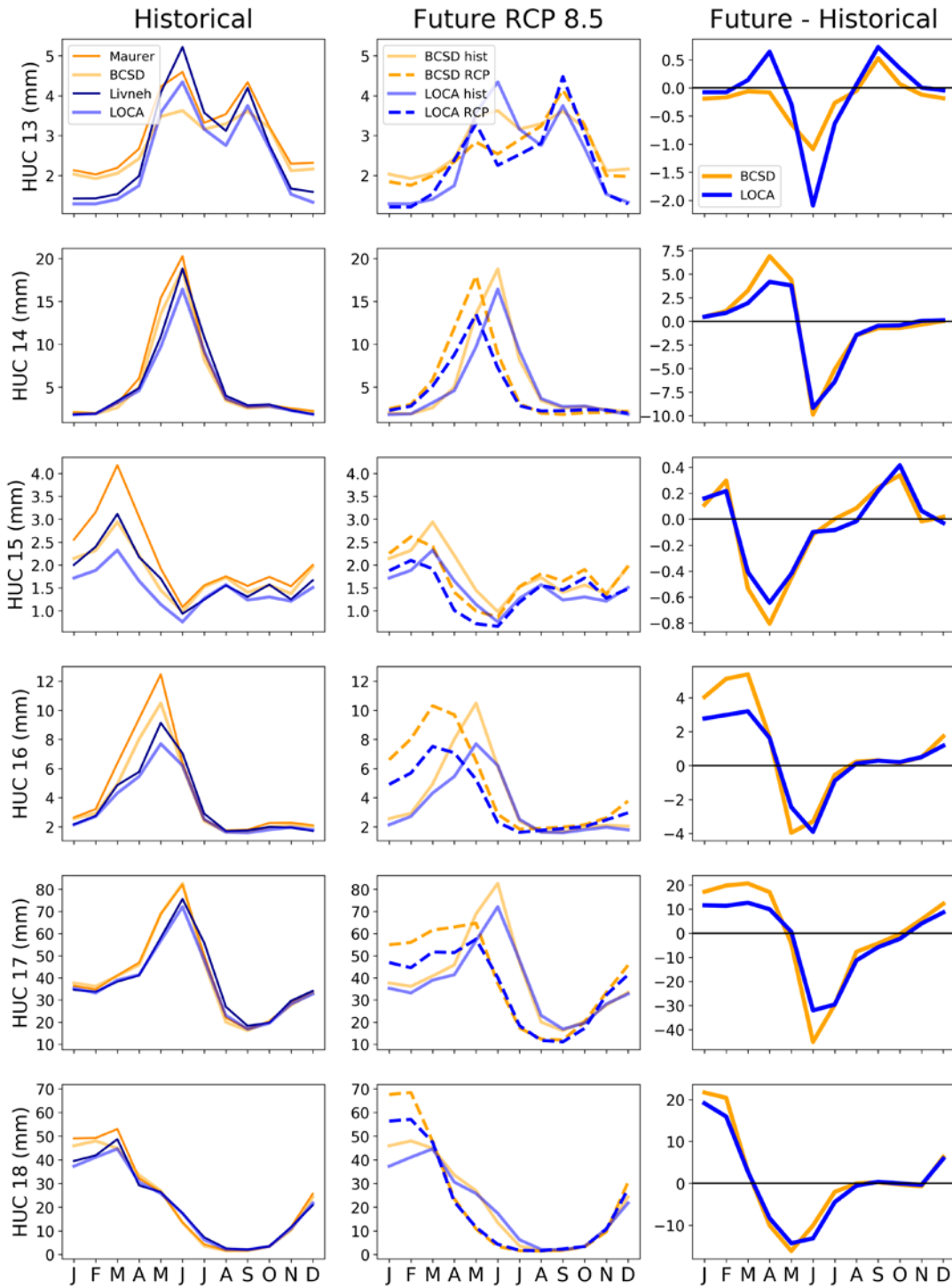
Figure D3. Historical and future basin runoff for HUC2 13-18. Similar to Figure 16 but for but HUC2 13-18. Values are in mm, averaged across the watershed area.

Appendix E. Daily Streamflow Exceedance Probabilities in 43 basins in the Western CONUS
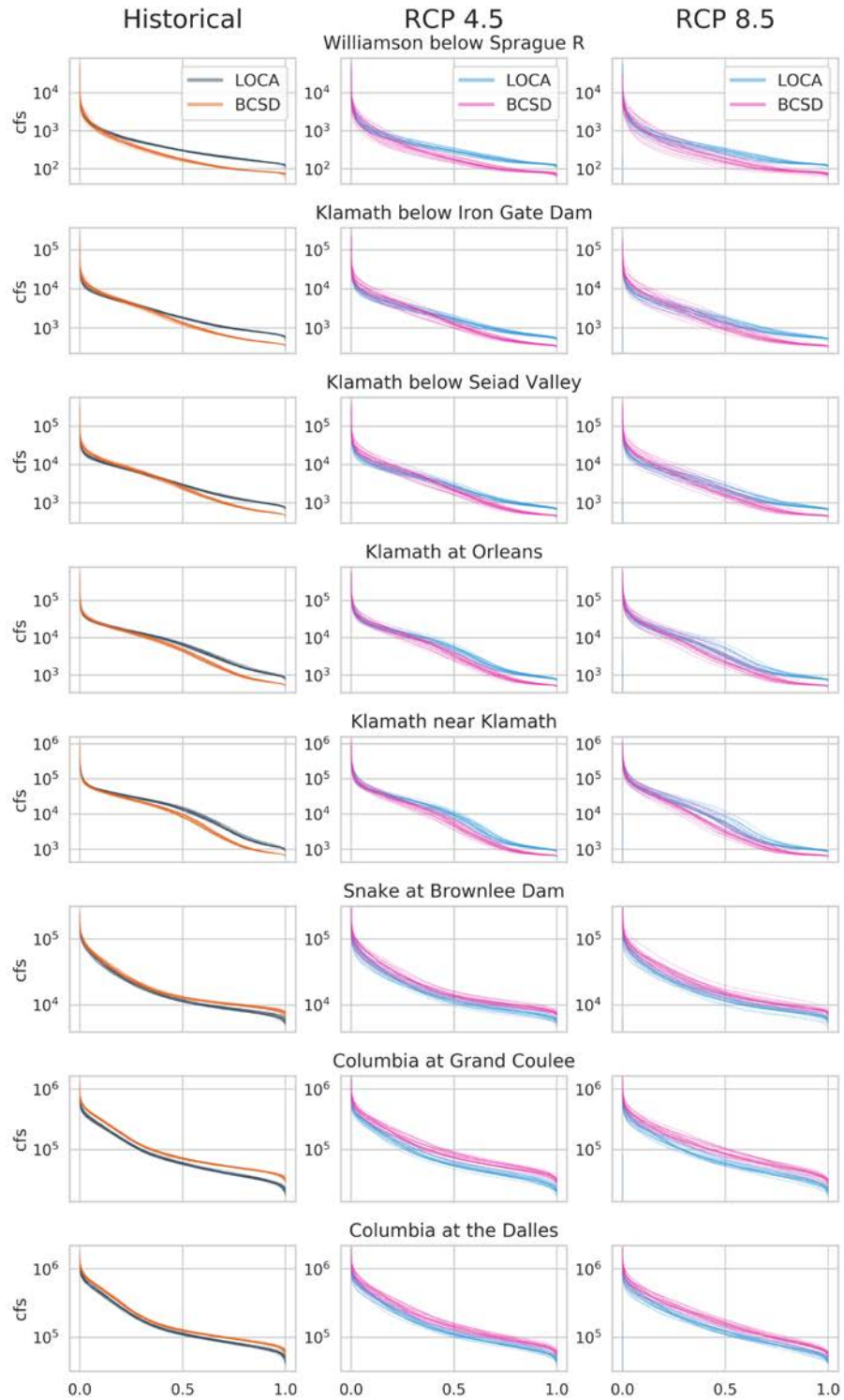
Figure E1. Locations 1-8

Exceedance probability
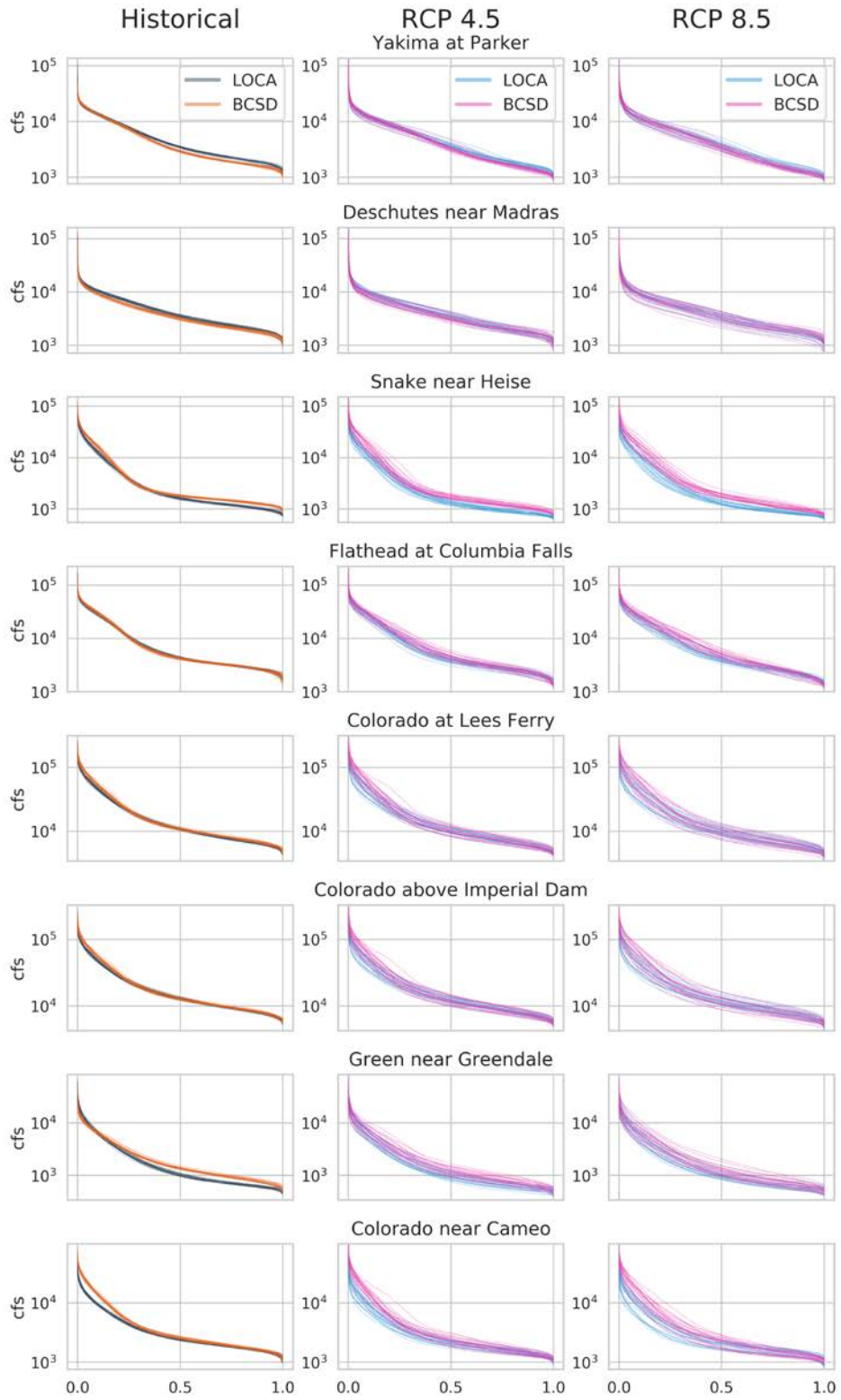
Figure E2. Locations 9-16

Exceedance probability
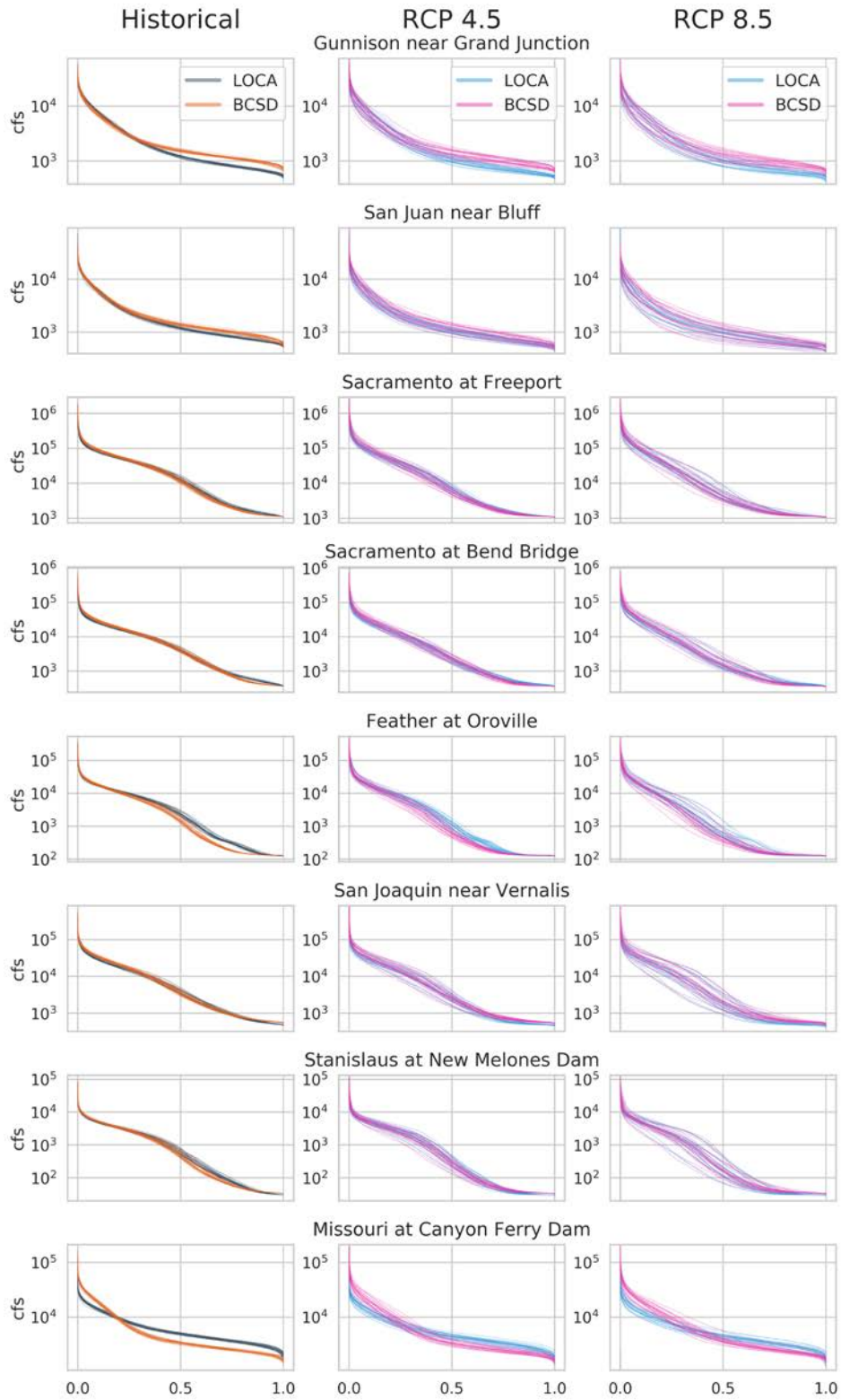
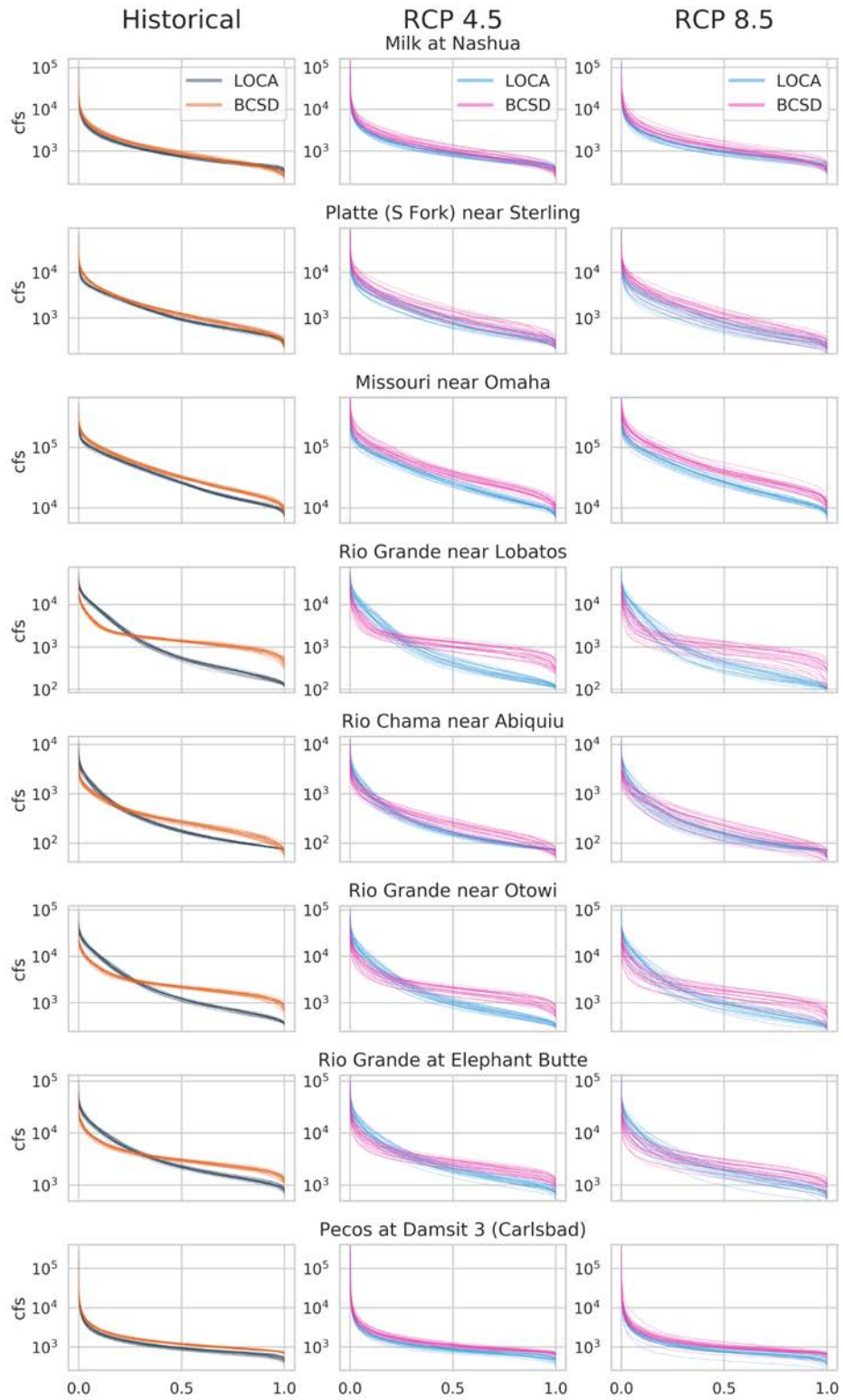**Figure E3. Locations 17-24**

Exceedance probability

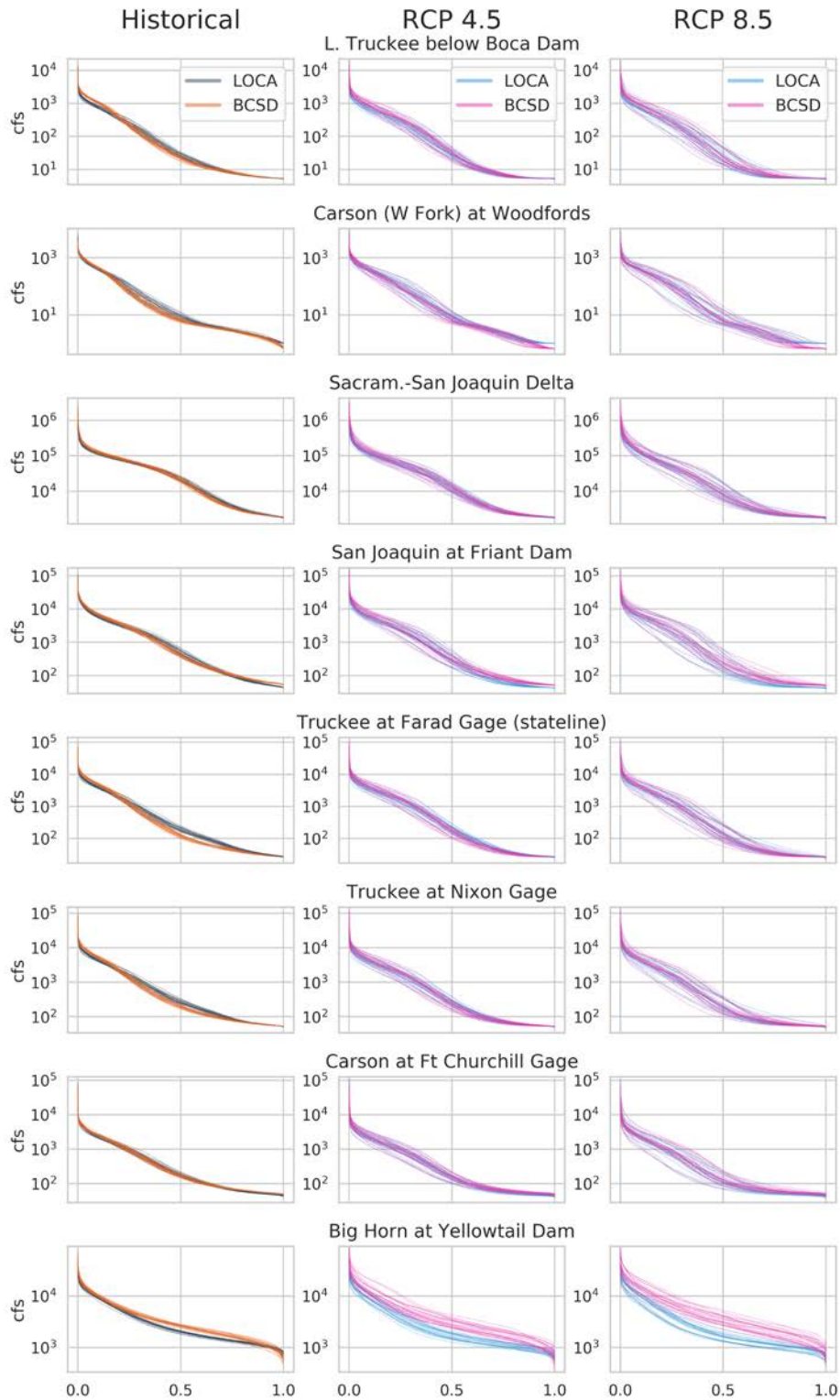**Figure E4. Locations 25-32**

Exceedance probability

Figure E5. Locations 33-40
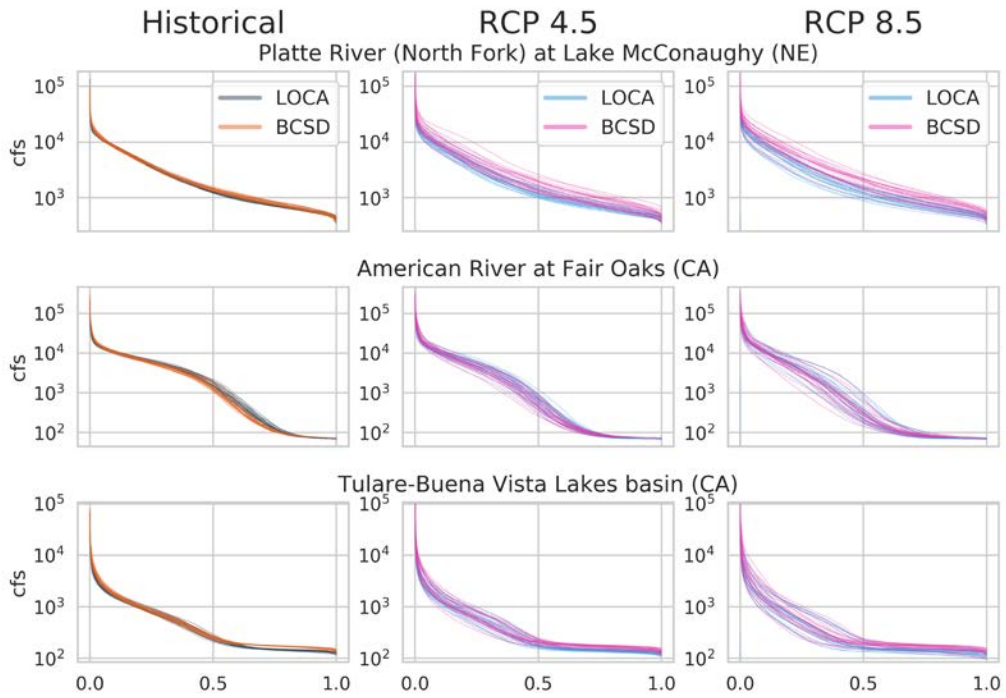
Exceedance probability

Figure E6. Locations 41-43

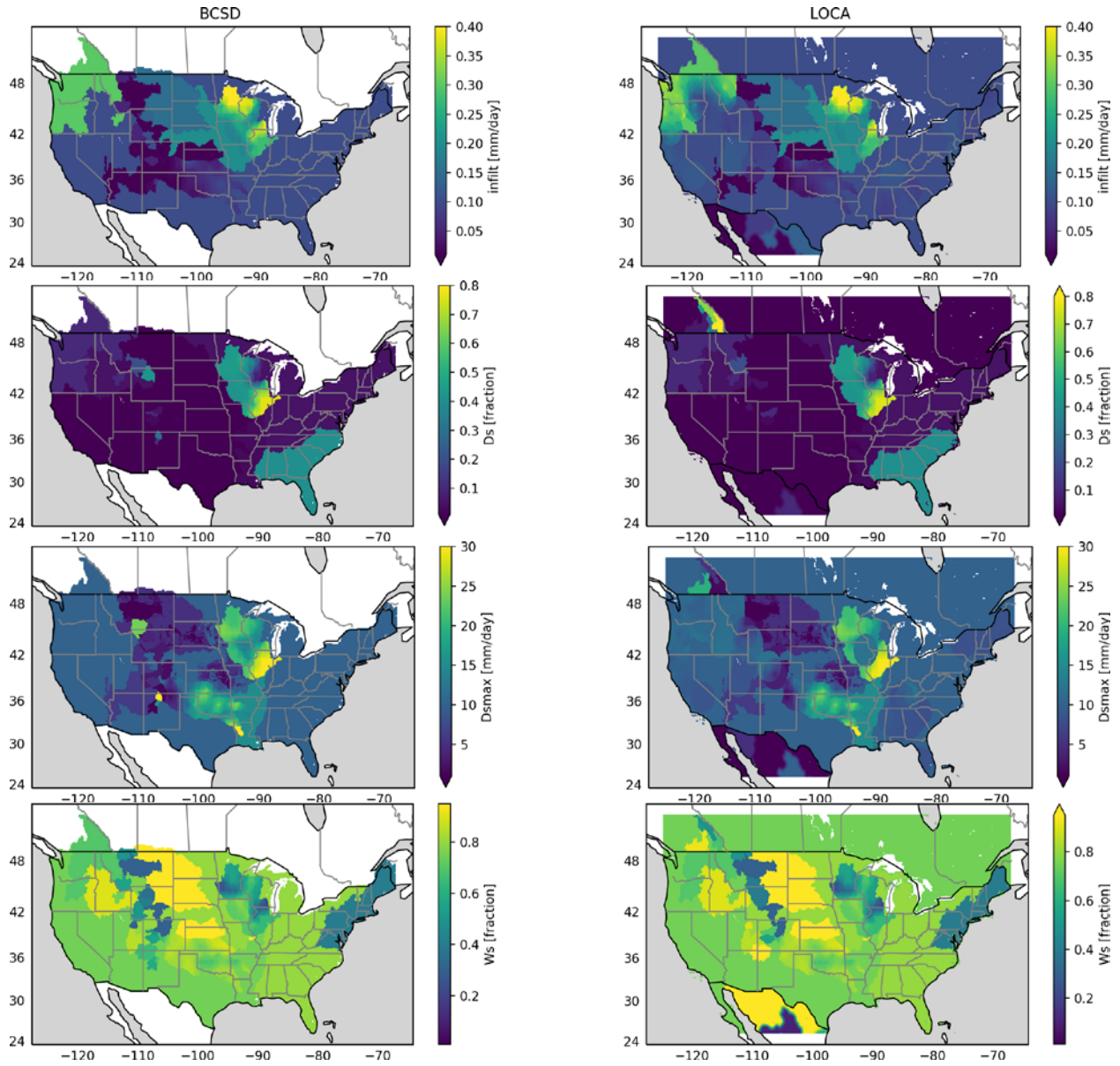# Appendix F. Maps of VIC parameters used in BCSD-VIC and LOCA-VIC.

Figure F1. Maps of 4 frequently calibrated VIC parameters from BCSD (left) and LOCA (right), infilt, Ds, Dsmax, and Ws