

## Peer Review Report

### Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections

Date:

June 23, 2020

Originating office:

Bureau of Reclamation, Research and Development Office, in coordination with US Army Corps of Engineers Climate Preparedness and Resilience Program

Reclamation roles:

Director or delegated manager: Levi Brekke, Chief, Research and Development Office, Bureau of Reclamation

Peer Review Lead: Kenneth Nowak, Water Availability Research Coordinator, Research and Development Office, Bureau of Reclamation

Peer Review Scope:

The aspect of the tech memo that was subject to peer review is the structure of the comparison between BCSD CMIP5 VIC and LOCA CMIP5 VIC simulations. The LOCA and BCSD downscaling methods were not the subject to this peer review; they are sufficiently reviewed and documented in peer-reviewed scientific publications. Further, datasets used in those downscaling methods are also not themselves the subject to this peer review; they also have been sufficiently reviewed and documented in peer-reviewed scientific publications.

Peer reviewers were asked to provide responses relative to the following questions:

Question 1. To what extent do the evaluations in the tech memo provide an appropriate high-level characterization for the purpose of describing differences in outputs between the datasets?

Question 2. How clear and convincing is the conclusion that many of the differences between outputs in the two datasets can be traced back to the different observational training data sets used for the different downscaling techniques?

Peer Reviewers

Stacey Archfield, PhD

Research Hydrologist, U.S. Geological Survey, Reston, VA

Daniel Broman, PhD, PE

Hydrologic Engineer, Bureau of Reclamation Technical Service Center, Denver, CO

Christopher Frans, PhD, PE

Civil Engineer, U.S. Army Corps of Engineers, Seattle District, Seattle, WA

## Summary of Reviewer Comments

Regarding two above specified questions in the peer review scope:

Reviewer 1 noted:

“The evaluation does a thorough job presenting high-level characterization of differences between LOCA- and BCSD- derived products.”

“Some clarification is needed in other differences between BCSD-VIC and LOCA-VIC, namely on the parameter sets, however the report is convincing in saying that these differences are minor compared to the historical training dataset, downscaling bias-correction method, or downscaling method apart from bias-correction.”

Reviewer 2 noted:

“The technical memo provides a thorough analysis and description of differences between the two datasets.”

“Aside from comments provided for question 2 below, dataset descriptions, the analyses, and discussions included in the document fulfill the purpose of describing differences in outputs between the datasets.”

Reviewer 3 noted:

“Overall, the observations and conclusions are well-supported and justified. This represents a significant contribution that will be of immediate use to the community. The results are communicated clearly and concisely in many places, particularly the conclusion section, where the main findings are clear and nicely summarized.”

Other feedback from reviewers included requests for additional methods detail, a range of typos/clerical corrections, minor clarifications, recommended references, and suggestions for improving results presentation. Except for a small number of requests that were beyond the scope of this report, all suggestions/feedback were accepted/addressed. Full details are available in below comments disposition table.

**LOCA Hydro Tech Memo**

Reviewer Name	Page(s) or other reference location	Line Number(s) if applicable	Priority	Comment	Response
Reviewer 1	page 2		Clerical	Abbreviations and Acronyms -add hyphen between Bias and Correction and change Equivalents to Equivalent	Done
Reviewer 1	page 3		Clerical	add DOI to Maurer et al. (2007) reference add link to Reclamation (2011) reference	Done
Reviewer 1	page 4		Clerical	move year in Pierce et al. (2014) reference to after authors rather than end, add DOI, update the DOI format in Pierce et al. (2015) to doi: convention	Done
Reviewer 1	Page 5	5th	Clerical	change contiguous to conterminous	Done
Reviewer 1	Page 5	8th	Substantive	add in description of runoff, either runoff (total runoff) like the introduction or runoff (runoff + baseflow) like later sections of the report	Done
Reviewer 1	Page 5	8th	Clerical	change equivalences to equivalence	Done
Reviewer 1	Page 5	17th	Clerical	Wide range of their individual requirements sounds odd maybe remove wide range?	agreed, we have changed it to "based on all of their requirements."
Reviewer 1	Page 5		Substantive	The introduction distinguishes between hydrologic and meteorological variables, with precipitation listed as a meteorological variable. However, in the 28th line the report states Differences in hydrologic variables ... LOCA generally has less change in average precipitation ... Maybe this distinction doesn't matter here, but for clarity I thought I'd point it out.	We agree that the wording was strange. The focus of this sentence was on the last section (that ET and runoff have less change) so we have re-arranged the phrasing in the sentence to be more consistent with likely reader expectations.
Reviewer 1	Page 5		Clerical	Wood et al. reference should be 2004 not 2014	Fixed
Reviewer 1	Page 8		Substantive	hydrologic model is the term that is used consistently in the report. in the 3rd line it states VIC hydrology model. I know this is a common way of referring to the VIC model, but for consistency it may be good to change hydrology to hydrologic.	Changed here and elsewhere "hydrology model" was used
Reviewer 1	Page 8	33rd	Clerical	Wood et al. 2004 reference is repeated twice	Fixed
Reviewer 1	Page 8	35th	Clerical	define CONUS here	Done
Reviewer 1	Page 8	46th	Clerical	I believe the reference for the historical Maurer dataset should be Maurer et al. 2002 not Maurer et al. 2007	Fixed
Reviewer 1	Page 9	7th	Clerical	7th line says all 23 GCMs but these have not yet been defined or discussed. Suggest removing the all or adding in a brief description of this subset above.	removed the "all 23" because this will be described later
Reviewer 1	Page 9	18th	Clerical	change daily value to daily values, Swap order NCA and CA State Assessment?	Done and done
Reviewer 1	Page 10	8th	Clerical	I believe the reference for the historical Maurer dataset should be Maurer et al. 2002 not Maurer et al. 2007	Fixed
Reviewer 1	Page 10	26th	Clerical	sentence starting with BCSD ... runs on. Split into two.	Done

Reviewer 1	Page 10	28th	Clerical	is the ca. needed?	It may not be needed, but there have periodically been updates to such datasets, as a result, it is possible that some portions of the data were only released in 2011, or that some model runs were completed and released early through institutional portals in 2009.
Reviewer 1	Page 11	last line	Clerical	last line – focusing on broad hydro-climatological features; other differences may be apparent at different scales. What is meant by broad?	Large scale, we have modified the wording.
Reviewer 1	Page 12	2nd paragraph	Clerical	see comment below about term definition convention. CONUS is already defined on p8.	Fixed
Reviewer 1	Page 12	5th paragrph	Clerical	terms already defined (ET, runoff, SWE). Not sure what convention is being used here. If the convention is to redefine each section (1, 2, 3 ...) then ok, otherwise remove. If remaining, replace equivalents with equivalent.	Fixed, the general intent was to define them once in the document, though the executive summary is treated separately.
Reviewer 1	Page 14	1st paragraph	Clerical	... and BCSD having greater declines in the western U.S., especially in the Canadian portion of the Columbia. As the Canadian portion of the Columbia is not part of the western U.S., suggest adding an and before especially.	We have changed western U.S. to western portion of the domain
Reviewer 1	Page 16	Fig 2 / Fig 3	Clerical	Figure 2 is smaller than Figure 3 because of the longer caption. Make the same size?	Done
Reviewer 1	Page 17	2nd	Clerical	change comparison of to comparison to	Done
Reviewer 1	Page 17		Clerical	observational dataset is used interchangeably with gridded observed dataset on this page and further in the report. suggest picking a consistent set of terms (e.g. gridded observed dataset and observed dataset) 'Forcing dataset' is also used (see below)	We have now tried to use "forcing dataset" only when we are referring to the data used to run vic (e.g. this could be obs, BCSD, or LOCA) and we have removed the word gridded from observed dataset references
Reviewer 1	Page 17	3rd paragraph	Clerical	The consistent differences between the LOCA and BCSD datasets appears as the fine-scale polka dot like feature surrounding individual weather stations ....This polka dot correspondence is also present in the precipitation dataset (Figure 5). The first sentence is referring to differences in temperature and this should be stated (and Figure 4 referenced), especially with the also in the second sentence introducing the precipitation differences.	We have rephrased and now reference figure 4 too.
Reviewer 1	Page 18	Figure 4	Clerical	Caption missing period after Figure 4	Fixed
Reviewer 1	Page 20	1st	Clerical	See comment above about defining variables. Variables are redefined again here	see above, fixed.

Reviewer 1	Page 20	1st	Clerical	runoff is defined here in more detail here. Should this remain, or be moved to an earlier section where the variables are first defined?	We now define it as (surface+baseflow) earlier and have removed this definition. We have also changed all references to "total runoff": to simply "runoff" since we have already defined that to be the term that encompasses both, and it could be confused with a long term summation over time.
Reviewer 1	Page 20	3rd	Clerical	not sure the convention for mentioning the VIC model. The Abbreviations and Acronyms page defines VIC as just Variable Infiltration Capacity rather than Variable Infiltration Capacity hydrologic model. as was used in previous reports. Suggest either changing the definition of VIC and keeping the it's stand-alone usage in the report or if using the current definition, using VIC model or VIC hydrologic model in the text.	This is constantly a source of confusion for all models, many papers I've reviewed refer to the VIC model, while others just call it VIC... same is true for WRF, etc. We will add "hydrologic model" to the definition to be consistent with past reports. I personally tend to just think of it as the name of the model rather thinking of it in terms of the words that name represents and that is probably as much a personal preference than anything else.
Reviewer 1	Page 20	2nd, 3rd, 4th Paragraphs	Clerical	observation dataset, and training dataset are all used interchangeably. A few consistent terms should be chosen.	We have reviewed usage, and tried to limit the use of "training dataset" though it is useful in a few cases to emphasize that that is all it is to the methods. In those cases we now refer to the training observation dataset
Reviewer 1	Page 20	4th paragraph	Clerical	add reference for MTCLIM algorithm	Done
Reviewer 1	Page 21	2nd paragraph	Clerical	change total runoff to runoff to match defined terms	Done
Reviewer 1	Page 21	3rd paragraph	Clerical	change Livneh and Maurer dataset to Livneh and Maurer datasets	Fixed
Reviewer 1	Page 21	3rd paragraph	Clerical	is Livneh et al. 2014 cited when the dataset is first referenced, or is it just Livneh 2015?	This is now changed from 2014 to 2013 (2014 was a correction) it is cited early on, but 2015 is the dataset that extends into Canada and Mexico and was used in the meteorology training for LOCA .
Reviewer 1	Page 21	3rd paragraph	Clerical	should PRISM be defined? Reference?	This only needed to refer to the "climate normal" used, so we have removed PRISM from this sentence.
Reviewer 1	Page 21	4th paragraph	Clerical	SWE already defined	fixed

Reviewer 1	Page 21	4th paragraph	Substantive	are the SWE differences significant?	We did not mean to imply statistical significance if that is what the reviewer means. For clarity, we have changed our use of the word "significant" so "substantial" here and throughout. No tests for statistical significance have been performed in this report as that is beyond the scope.
Reviewer 1	Page 21	4th paragraph	Substantive	the better resolved topography offers a good explanation for SWE differences seen in the western mountains, but is there an explanation for SWE differences seen in New England?	No, this is an interesting point that we didn't have a good explanation for so we did not go into detail on it. It is now described explicitly in this section as likely connected to the more subtle features of the meteorology. The difference in SWE in New England matches the differences in the Livneh and Maurer VIC simulations here. The two datasets have more differences than just topography, the increased degree of interpolation required in Livneh changes the statistics of the meteorology, as do small differences in stations used.
Reviewer 1	Page 24	1st paragraph	Clerical	the future climate changes are used to describe what is being evaluated, however, in the Introduction representation of projected changes ... is used, and in the section heading future climate conditions is used.	We have added the word projected around all descriptions of future climate to make it clear that these are all projected. We have left the headings, and some of the introduction untouched as these more generally refer to the concept of future rather than the specific projections being analyzed.
Reviewer 1	Page 25	2nd paragraph	Substantive	BCSD has smoother changes than LOCA, resulting in locations being hotter and others being drier throughout the country. 1) Figure 9 shows both hotter and cooler differences between LOCA and BCSD 2) Figure 10 does show a dry differences between LOCA and BCSD, and similar to how Figure 9 is referenced, Figure 10 should also be referenced.	I'm not sure what is meant here. The second paragraph is all about changes in temperature, and does not reference changes in precipitation (dry differences). The third paragraph is about the changes in precipitation and references figure 10 in four places.
Reviewer 1	Page 25	2nd paragraph	Clerical	change ACCESS1.0 to ACCESS1-0	Done
Reviewer 1	Page 25	2nd paragraph	Clerical	For example, both BCSD and LOCA show increased warming in MIROC5 (and ...) over the regions that are represented in the GCM with higher elevation terrain ... as compared to what?	As compared to the surrounding areas with lower elevation terrain. This has been rephrased.

Reviewer 1	Page 27	Figure 10	Clerical	Fig 9 should be changed to Fig. 9	All "Fig" references have been changed to "Figure" throughout the text for consistency
Reviewer 1	Page 27	2nd paragraph	Clerical	capitalize River in southern Mississippi river	Done
Reviewer 1	Page 28	1st paragraph	Clerical	capitalize River in the upper Columbia river basin	Done
Reviewer 1	Page 28	1st paragraph	Clerical	remove the before with the ACCESS1-0	Done
Reviewer 1	Page 29	1st paragraph	Substantive	...and possibly due to changes in the VIC parameters across the border. Were VIC parameter differences, either spatially, or between BCSD-VIC and LOCA-VIC, looked at and could they be attributed to differences in the modeled results?	This is something we wanted to do but ran out of time. We now have added appendix F to illustrate the similarities and differences. There are too many parameters to show everything, so we have plotted 4 of the more commonly calibrated parameters. We think this sort of documentation is an excellent addition, thank you.
Reviewer 1	Page 30	1st paragraph	Clerical	check usage of gridcell vs grid cell	Done
Reviewer 1	Page 30	1st paragraph	Substantive	he VIC model can represent sub-grid variability through the use of multiple elevation bands. It can, but what was done in this study? SWE differences are also seen in the Appalachians at higher elevations (WV, NY, New England), but there are slightly less negative, but still large differences in lower-elevation areas in New England. Are there thoughts on the cause of this?	Yes it was, though sub-grid snow bands still provide melt water into a single grid-cell wide soil column, so this is not a complete solution. We don't have a good reason for the cause of the noted differences.
Reviewer 1	Page 31	1st paragraph	Clerical	change approaches to methods	We have changed the word approaches throughout whenever referring specifically to BCSD and LOCA, but left it when writing more generally.
Reviewer 1	Page 31	1st paragraph	Clerical	change Figs. 2 and 3. to Figures 2 and 3 or Figure 2 and Figure 3.	Done
Reviewer 1	Page 31	1st paragraph	Clerical	climate change signal seen in the CMIP5 depending on ...Change to either seen in CMIP5 or seen in the CMIP5 dataset.	Fixed
Reviewer 1	Page 32	2nd paragraph	Clerical	italicize Mauran	Done
Reviewer 1	Page 32	figure 14	Clerical	change 2070-2100 to 2070-2099. This occurs in a few places.	Sorry, this occurred in two places where a different author was writing. The nomenclature depends on whether you consider 2100 to mean through Jan 1 2100 or Dec 31 2100 (i.e. inclusive or now). It is further confused by the fact that some GCMs provide the year 2100 and others do not. We have modified to 2099 to for consistency.

Reviewer 1	Page 33	1st paragraph	Clerical	when water arrives and is in our rivers and streams. 'Our' is a departure from the voice used in the report to this point, at least since the introduction.	Changed
Reviewer 1	Page 33	1st paragraph	Clerical	defining terms again. see comments above	Fixed
Reviewer 1	Page 33	1st paragraph	Clerical	HUC2 is used before, but to this point not referred to as HUC2 watersheds	Fixed
Reviewer 1	Page 33	1st paragraph	Clerical	we use, similar comment to Our above	Changed
Reviewer 1	Page 33	3rd paragraph	Clerical	downscaled-based VIC datasets; observational-based VIC dataset. Check against terminology used earlier in report	BCSD5-BCSD3 report did not draw as clear a distinction, and frequently referred to the "historical hydrology", though this leaves ambiguity between the historical GCM forced hydrology and the historical observation forced hydrology. We prefer the more explicit terminology in this report.
Reviewer 1	Page 33	4th paragraph	Clerical	adding descriptions to HUC labels, or adding names to Figure 15 could help the understanding especially for those without a good mental map of CONUS river systems.	This is a good point we feel that since Figure 15 is already a map, it reveals the context explicitly. We have added labels to Figure 16 to minimize the number of times the reader might have to scroll back up to Figure 15 to identify HUC numbers out of context.
Reviewer 1	Page 34		Clerical	Are RCP 4.5 results for the 18 HUC2 watersheds also found in Appendix D?	No, since the RCP45 changes are just a smaller change than the 8.5 it was not deemed necessary.
Reviewer 1	Page 34		Clerical	Text on p33 refers to HUC 1, 5, 9, 11, 13, 14, 15. HUC 3, 5, 9, 11, 14, 17 are shown in Figure 16. Should Figure 16 include the HUCs mentioned in the text, or was the rational to include a geographic spread on HUCs in this figure?	Yes, we merely wanted to cover a range and show a few key examples. Later text also refers to additional HUCs and we didn't want to fill the text with all of the HUCs
Reviewer 1	Page 35	1st paragraph	Clerical	The changes in BCSD-VIC and LOCA-VIC runoff across these same regions ...which same regions?	This was supposed to refer to the same regions that were just presented in Figure 16. We have now made that explicit, because I agree that in reading the text it is not clear.
Reviewer 1	Page 35	1st paragraph	Clerical	...and not changing the model parameters could result in a model calibrated for the statistics of the Maurer dataset while being driven my meteorology with the statistics of the Livneh dataset. 'could', what was actually done?	It is a little of both, we have clarified.
Reviewer 1	Page 35	2nd paragraph	Clerical	add abbreviation definitions to season definitions, e.g. winter (Dec-Jan-Feb; DJF), to correspond with figures (e.g. Figure 17).	Done
Reviewer 1	Page 35	2nd paragraph	Clerical	change percentage points to %.	Done
Reviewer 1	Page 36	Figure 17	Clerical	change 2100 to 2099	Done
Reviewer 1	Page 37	3rd paragraph	Clerical	change LOCA datasets to LOCA dataset	Done



Reviewer 1	Page 37	3rd paragraph	Clerical	change Columbia river basin to Columbia River basin	Done
Reviewer 1	Page 39	Figure 19	Clerical	change GCMS to GCMs	Done
Reviewer 1	Page 40	1st paragraph	Clerical	change both techniques to both methods	Done, both here and elsewhere in the document.
Reviewer 1	Page 40	1st paragraph	Substantive	The differences between ensemble mean values for LOCA and BCSD are generally smaller than the change signal, though locally they can be larger In the first clause, what geographic extent is the comparison being made across. From the second clause, it's larger than 'local', but is it across the entire US, regionally, ...?	We have clarified with examples of specific places in which the difference is greater than the ensemble mean change.
Reviewer 1	Page 41	2nd paragraph	Clerical	change downscaled products to datasets	Done here and elsewhere
Reviewer 1	Page 41	2nd paragraph	Clerical	change Fig. to Figure	Done
Reviewer 1	Page 43	1st paragraph	Clerical	snow melt and snowmelt are both used. Choose one and use consistently	Done
Reviewer 1	Page 44	1st paragraph	Substantive	The ensemble mean changes (top row), however, show significantly larger increases in the central and western U.S. in both LOCA-VIC and BCSD-VIC ... larger increases compared to what?	We have specified that is is larger than the CONUS average, and rather than making it sound like it might be everywhere across the western and central US, we have clarified that it is only in a few regions.
Reviewer 1	Page 45	1st paragraph	Substantive	These daily flow values have not been bias corrected and these basin are not consistently calibrated. Does this mean that some basins were calibrated? I think this needs some more explanation.	This is a very confusing aspect of continental scale VIC applications and it deserves more discussion. We have added additional discussion here. The quick answer is that we don't have observed flows everywhere, so CONUS wide traditional calibration is impossible, and the calibration that has been performed in VIC has been an ad hoc process done by dozens of (mostly grad students) over a decade or more. So some specific basins in the domain have been calibrated (likely including one or more of the basins in question here) but there is not a catalog of calibrations that have been performed. Many regions have only default parameters that are based on, for example, pedo-transfer functions that we know do not adequately describe the sub-surface processes.
Reviewer 1	Page 46	2nd paragraph	Clerical	extra space before Generally	Fixed
Reviewer 1	Page 46	2nd paragraph	Clerical	Figure 26 shows seven locations not needed. see Appendix E for all 43 add basin.	Fixed

Reviewer 1	Page 47	1st paragraph	Clerical	such as these is awkward	Appendix figures have been labeled with a letter for the appendix followed by the number within the appendix; however, this should have been to Figure A3. We have also updated "Fig." to "Figure"
Reviewer 1	Page 47	1st paragraph	Clerical	Formatting notes – 1) change Williamson R to Williamson River; 2) extra space between Fa and lls in 12; 3) Damsit should be Damsite in 32?; 4) atYellowtail missing space in 40;	Table updated.
Reviewer 1	Page 49	1st paragraph	Clerical	increase in the frequency of extreme daily runoff events. Extreme daily runoff events not defined.	This was defined later, but we have now defined it more explicitly here.
Reviewer 1	Page 49	1st paragraph	Clerical	both gridcells and grid cell used. Pick one and use consistently	Done
Reviewer 1	Page 50	1st paragraph	Clerical	change techniques to methods	Done
Reviewer 1	Page 50	1st paragraph	Clerical	see gridcell/ grid cell comment above	Done
Reviewer 1	Page 50	1st paragraph	Clerical	it inherits more of the GCM length scale I'm not sure if the concept of GCM length scale is common and this could be confusing	We have rephased as "spatial autocorrelation distance", though I sort of like length scale more.
Reviewer 1	Page 50	2nd paragraph	Clerical	Forcing data. See comments above about terminology for the Livneh and Maurer datasets	Here forcing data could refer to whichever dataset is used to force VIC, e.g. observations or downscaled data, so I believe it is more appropriate to use.
Reviewer 1	Page 51	1st paragraph	Clerical	move MTCLIM reference to it's first use in the report	It has been added to the first reference and kept here for easy reference.
Reviewer 1	Page 51	2nd paragraph	Clerical	observation-based datasets. See comments above about terminology for the Livneh and Maurer datasets	This has been revised to be more consistent
Reviewer 1	Page 51	2nd paragraph	Clerical	...as the observed dataset... add s to dataset	Done
Reviewer 1	Page 56	1st paragraph	Substantive	...so river basins and local watersheds can be modeled to provide as good of an approximation of the real hydrologic system as possible. It doesn't seem to me that this is the main point of this effort. A reasonable representation of the real hydrologic system is needed, but the main goal is to provide reasonable representations of the real hydrologic system with consistent historical and future projection data to examine changes. If as good of an approximation of the real hydrologic system as possible is the real goal, models should be calibrated and verified against observations.	Well put, we have modified the first paragraph in the conclusion to make this point.
Reviewer 1	Page 56	2nd paragraph	Substantive	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. Didn't selecting the subset of 23 models remove the effect of the different combination of GCM/RCP?	Here we were referring to the more general concept, which is why we note that they have less effect here. We now note that explicitly.

Reviewer 1	page 57	3rd paragraph	Substantive	If a GCM does not represent important aspects of the regional climate ...users are encouraged to explore the physical processes behind any changes projected to understand and contextualize the reliability of those projections. Are the data accessible to do this? Should the report reference the ESGF website? Could users navigate this site? It appears that regridded raw GCM data (1deg) may be available on the GDO website. If so, could the report highlight this and point users to these resources?	This goes beyond what the GDO site supplies, and while we mention the CMIP archive at the beginning, the purpose of this paragraph is to draw back from the technical components to encourage users to think about the hydro-climate in their basin and whether or not the processes discussed here are sufficient for their purposes.
Reviewer 1	Page 66	Figure caption	Clerical	refers to Fig. F. Should 'F' be replaced with a number? Also see comment about consistency in referring to figures.	Appendix figures have been labeled with a letter for the appendix followed by the number within the appendix; however, this should have been to Figure A3. We have also updated "Fig." to "Figure"
Reviewer 1	Page 68	Figure caption	Clerical	refers to Fig. F. Should 'F' be replaced with a number? Also see comment about consistency in referring to figures.	see above
Reviewer 1	Page 69	Figure title	Clerical	Figure title – Snow Water Equivalent should not have an 's' at the end	Fixed
Reviewer 1	Page 72	Figure title	Clerical	RCP is mislabeled. Should be RCP 4.5 rather than RCP 8.5	Done

Reviewer 1	References	References	Clerical	<p>Add DOI to Fowler et al. 2007  Add year to Gao et al.  Change link in Haerter et al. 2011 to DOI  Add periods after initials in Harding et al. 2012  Add DOI to Livneh et al. 2014  Add DOI to Livneh et al. 2015  Add link to Lukas et al. 2014 ...  <a href="https://wwa.colorado.edu/climate/co2014report/Climate_Change_CO_Report_2014_FINAL.pdf">https://wwa.colorado.edu/climate/co2014report/Climate_Change_CO_Report_2014_FINAL.pdf</a>  Remove https:// from DOI in Mauran 2013  Add DOI to Mauran 2016  Add DOI to Maurer 2007  Maurer 2014 – replace period before DOI with comma, format DOI using doi: convention  Add periods after initials in Mendoza et al 2015; remove ‘a’ after 2015  Remove https:// from DOI in Michelangeli et al. 2009  Remove https:// from DOI in Mizukami et al 2016 format DOI using doi: convention  Add DOI to Pierce et al. 2014  Pierce et al. 2015 – replace period with comma before DOI, format DOI using doi: convention  Add link to Pierce et al 2018  Add link to Reclamation 2011; replace comma after year with period  Move Thornton and Running 1999 reference after ‘S’ references; format DOI using doi: convention  Switanek et al. 2017 – missing journal name, DOI  Sankarasubramanian et al. 2001 – add DOI  Wilby et al. 2004 – add missing periods after initials, add link?  Wood et al. 2004 – add DOI</p>	Done
Reviewer 1	General	Figure title	Clerical	<p>In figure captions, previous figure, Fig. XX, and Figure XX are used interchangeably. I suggest using Figure XX to refer to other figures, matching the convention found in the text.</p>	Done
Reviewer 1	General	introduction	Substantive	<p>There is no section in the report that describes the study area, aside from CONUS. The fact that the Canadian portions of the Columbia and Milk River (and perhaps others) are evident from Figure 1, but it could serve the report well to mention this. I’m imagining the reason that the Mexican portions of the Rio Grande (and others) are missing are that BCSD does not extend south of the US border, but LOCA does. It could also serve the report to mention this.</p>	Good point. This is now noted early on in the general discussion of BCSD and LOCA.
Reviewer 1	General	introduction	Clerical	<p>Similar comment as the study area for the time periods used for comparison. Figure 1 identifies the historical period as 1970-1999, and Figure 2 identifies the future period as 2070-2099, but these time periods are not defined in the text.</p>	Good point, we have added this to section 3.1

Reviewer 1	General	Throughout	Substantive	Terminology for historical VIC simulations should be consistent, while both 'observed' and 'modeled historical' are used throughout the report.	We now define this more explicitly. "observed" refers to VIC simulations run with observed meteorology, while modeled historical refers to simulations using the GCM historical data.
Reviewer 1	General	Throughout	Clerical	Re: geographic terminology use consistency: there are locations where U.S. is used and locations where the CONUS is used for example and one term should suffice. In addition the 'river' in 'river basin' is sometimes capitalized and sometimes not and a consistent usage (I suggest capitalized) should be used	Have changed all named rivers to uppercase "R"s and changed references to US to CONUS.
Reviewer 1	General	Throughout	Clerical	Corresponding figures, for example D1-D3, or E1-E4, are not consistently aligned on the page. For the first figure with a caption this makes sense, but for the latter figures, there isn't consistent placement of the figure on the page. Ideally I'd suggest trying to make all figures the same regardless of whether it has a longer caption (the first figure), but I can see why making the figures fill the page is nice. At the very least the latter figures should align.	We have made these figures a common width and centered on the page.
Reviewer 1	Peer review plan Q1			The evaluation does a thorough job presenting high-level characterization of differences between LOCA- and BCSD-derived products. I particularly liked the organization of the report, which first presents high-level findings, and then discusses changes in each of the variables in an order corresponding to the steps of modeling. These changes are also discussed at first broadly (historical and future averages) before looking at shorter timescales and behavior of the extremes. Apart from some clarity to the geographic extent and time periods evaluated, I thought the background information was sufficient to understand the remainder of the report, and that appropriate references were provided for those looking for additional information. The appendices provide a good reference for users looking for information specific to their region. A thought would be would it be possible to provide these information in a digital format as well, either as tables, or a map, where the information could be queried	Thanks, we have also thought about trying to make more of this available digitally. We will discuss internally and see if we can at least make the python scripts that were used to generate all of these metrics available, and maybe we can find a place to host the summary metrics, as I agree sometimes that would make it easier for others to process. Building, e.g., a web interface to all of this is beyond current scope, though it would be pretty cool, and not all that hard if the data are in a cloud provider with a jupyter notebook style interface.

Reviewer 2	Section 3.2.2	Throughout	Substantive	<p>It is not clear how the hydrological model parameters used for VIC vary between LOCA and BCSD and where they came from. Potential differences in parameters are briefly noted in section 3.2.2 and 3.3. While this may be provided in the cited reports in sections 2.1 and 2.2, a brief explicit description of the parameters used would be valuable in section 2.1 and 2.2</p>	<p>The short answer is that there are too many parameters to document all of them. The parameters used in the LOCA runs were those derived for the Livneh dataset (though very little changes from the Maurer dataset). The parameters used in the BCSD runs were those from the Maurer climatology, though undoubtedly there have been a few modifications to them over time. These parameters should probably be documented in a publication sometime, but they are not now.</p>
Reviewer 2	page 57	First paragraph	Substantive	<p>On the first paragraph of page 57 “when the LOCA data are used, it is important to recalibrate VIC (or other models) to be consistent with the LOCA climate” Does this imply that the parameters used for LOCA were not calibrated for LOCA climate? If they weren’t calibrated for LOCA and are not the same as BCSD, what is the source of parameters?</p>	<p>These parameters come from the Livneh et al paper, and they were only partially recalibrated to match the VIC output from the Maurer dataset itself. This recalibration was performed for only a few grid points and only for a few parameters, which were then interpolated in space. See the circular features in the new appendix F showing in inverse distance function modifying parameters that otherwise match the BCSD parameters almost exactly.</p>
Reviewer 2	Page 46		Substantive	<p>On page 46 the text reports “This provides an identical network setup and routing algorithm and thus allows a direct comparison between the two datasets.” However, it seems that differences in flow duration curves will also be greatly influenced based on which hydrological parameters were used, perhaps more so than influences of routing.</p>	<p>Agreed, this was just to mean that the changes in the hydrologic model projections can be compared without worrying about differences in the routing too.</p>

Reviewer 2		Figure 3	Substantive	The divergent scale for temperature change makes it impossible to see spatial patterns in temperature change signal	It is not so much the divergence as the recent push for perceptually smooth color scales which make it harder to see as many gradients, but they make the interpretation of those gradients more accurate. We struggled a lot in this selecting colorscales, my biggest complaint is probably that the color range shifts so much between plots that to select a consistent range that covers all plots makes it hard to see changed on some individual plots. However, this emphasizes the point that other differences (e.g. between RCPs) are probably more important than the differences between BCSD and LOCA. Which is also one of the main points here. see <a href="https://www.osti.gov/servlets/purl/1338147">https://www.osti.gov/servlets/purl/1338147</a> <a href="https://www.sciencemag.org/news/2018/08/its-time-use-fewer-colors-scientific-figures-some-experts-say">https://www.sciencemag.org/news/2018/08/its-time-use-fewer-colors-scientific-figures-some-experts-say</a>
Reviewer 2	Page 30	last paragraph	Substantive	“This higher spatial resolution permits 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. However, the VIC model can represent sub-grid variability through the use of multiple elevation bands within a gridcell, so the link is not as direct as it would be in other hydrology models.” Isn't it more likely attributed to the Livneh training having colder temperatures at high elevations, rather than resolution?	The primary reason the Livneh dataset has colder mountain tops is its higher resolution, so yes, but the two are largely the same thing.
Reviewer 2		Figure 22	Clerical	No titles on top panels	Fixed

Reviewer 2	Page 43		Substantive	<p>“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 the hydrograph in snow melt dominated regions (Figure 16), with less peak seasonal streamflow in LOCA-VIC than in BCSD-VIC related to the higher elevation, colder snowpack and snowmelt occurring in the model when vegetation is more actively transpiring.” I don’t fully understand the description, would it be more simple to attribute this to the LOCA dataset having less precipitation and colder temperatures in the western mountains than BCSD?</p>	<p>We have tried to clarify this, as we agree that it wasn't written well. Simply having more annual precipitation does not necessarily correspond to greater peak streamflow amounts. In this region, the peak streamflow is often related to how rapidly snow can melt, and having higher/colder mountain tops leads to a later melt season, can spread the melt out over a longer time period, and means that more of it will come later in the year when potential ET is higher (leading to less water making it into the stream channel).</p>
Reviewer 2	Page 49		Clerical	<p>Rearrange text for clarity in reading, move description of figure 27 before discussion of figure 27</p>	<p>Good point, we have moved the descriptive sentence up in the text.</p>
Reviewer 2	Section 3.3.2-4		Comment	<p>I found section 3.3.2-4 really interesting I am glad you included these comparisons</p>	<p>Thanks!</p>
Reviewer 2	Page 57	Second paragraph	Substantive	<p>As I understand this step change is a statistical artifact of the downscaling methods, constraints were applied to the historical period and not the future. It would be valuable to comment on how this step change could influence period to period comparisons and what considerations could be made in lieu of it.</p>	<p>We have added a note describing how this was handled in a recent publication by Wobus et al.</p>
Reviewer 2	Peer review plan Q2		Substantive	<p>The memo provides convincing evidence that the observational historical training datasets explain many of the differences between LOCA and BCSD projections. This is a key challenge being grappled with in the use of several datasets in the applications community in the western United States. Given its importance, and being that this is key feature of the technical memo, I feel that there is opportunity to provide further support for this information:</p> <p>Use recent literature to further support the conclusion of the importance of observational based training sets on hydroclimate projections. The following two studies have looked at similar datasets in the Western United States. These use similar or the same statistical downscaling methods and training datasets.</p> <p>Alder, J. R., &amp; 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. <i>Water Resources Research</i>, 55, 2279–2300. <a href="https://doi.org/10.1029/2018WR023458">https://doi.org/10.1029/2018WR023458</a></p> <p>Jiang, Y., J.B. Kim, C.J. Still, B.K. Kearns, J.D. Kline, and P.G. Cunningham, 2018: Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. <i>Scientific Data</i>, 5, 180016. DOI: 10.1038/sdata.2018.16.</p>	<p>We have added these references, thank you.</p>



Reviewer 2	Peer review plan Q2		Substantive	<p>It would be valuable to provide a summary description of the methods used to generate each historical observational dataset in section 2, highlighting how they are different. Fragments of this is available in later discussions of LOCA-BCSD differences. However, having this information available in a central location in section 2 would be valuable when reading and interpreting section 3 content.</p> <p>The most familiar historical dataset in the hydrological applications community is PRISM. Providing a comparison of Livneh and Maurer to PRISM may further contextualize these differences in training datasets and be of use to the end user. However, this level of analysis on the historical datasets alone may be outside the scope of the memo and provide less value for the sole purpose of describing the differences in projections.</p>	Yes, we debated how much detail to add, and ultimately decided that level of detail was too much to put in here. The reason the observation datasets are different is probably not as important as the fact that they are different.
Reviewer 3	Section 2.3		Clerical	Consider a header for each of the first three bullets titled: "Timesteps", "Analog matching", and "Bias Correction Methods". This would direct the reader to key topics discussed in each bullet.	This is a really nice suggestion, we have added a short header to each of the bullets in this section now.
Reviewer 3	Overall		Clerical	There seems to be inconsistency in the presentation of the figures throughout the memo and some figures need to be checked. I would suggest going through the figures to make sure they are stand-alone and consistent in formatting across the memo	We have endeavoured to make most figures extremely consistent; however, there are a few figures that were generated using a different plotting program by David Pierce and we didn't want to either, 1) ask him to change his program, or 2) duplicate the effort, we have attempted to make the figures similar anyway, but recognize that this is not ideal. We have made the colorscales use consistent colors, but the two programs end up with subtly different real colors.
Reviewer 3		Figure 1	Clerical	Color schemes change across figures and within figures (see Figure 1 as an example) but there are examples throughout.	This is done deliberately to make it easier to distinguish between temperature (uses viridis for absolutes and BlueRed for deltas) and water related variables like ET, runoff, etc that use a bluegreen colorscale for absolutes, which is similar to the positive side of the BrownBluegreen colorscale that is used for deltas. In most places, the colorscales are quantized to make it easier to distinguish specific numbers from them, while in a few cases it is left continuous because values often fell within a single quantile.
Reviewer 3		Figure 9	Clerical	The title for Figure 9 gives the units of "mm" for temperature, which seems incorrect?	Done

Reviewer 3		Figure 10	Clerical	In Figure 10, the legend has changes are from -96mm to +96mm. Would it be more informative to report percent change? Later (in figure 19) you use percent change rather than absolute change.	We definitely debated this point, and have made the figures both ways. We ended up using absolute values instead of percent because in large basins (e.g. the Upper Colorado River basin) the absolute change effects the runoff more. For example, much of the Colorado river basin has very little precipitation, thus a small change would look like a larger percent change, and might appear more important than small percent changes in the mountains that would have a larger effect on water resources. In other basins, percent change might be more important of course.
Reviewer 3		Figure 17	Clerical	Figure 17 looks very different from the other figures and how they are presented? Is there a reason that the colors are different here than in other figures? Also in Figure 17, the seasonal abbreviations in the figure captions need to be added so that users can relate this to the abbreviated season names in the figure.	This is one of the figures made with a different plotting software package. We have used the "same" colorscale, but different packages produced slightly different actual colors. We have added the seasonal abbreviations.
Reviewer 3		Figure 22	Clerical	Figure 22 is missing the headings above each column of figures.	fixed, thanks.
Reviewer 3	Section 2.3			the last bullet has a number of references for which the reader can refer to supporting materials; however, in the earlier bullets, it would be helpful to provide citations to the statements of fact that are made in the first 3 bullets on p. 10.	While I agree in principle, the first three bullets are largely restatements of the algorithms used in the two methods, and the references for those are simply the method data papers, which are cited immediately above. The bullet on bias correction differences makes some slightly more contentious statements that have been argued about recently in the literature, so we felt it more important to add additional citations to this section.

Reviewer 3	Overall			<p>Generally a figure is introduced in the first sentence of a section and then all statements after that do not include any figure references. This is very difficult for the reader to understand where the supporting evidence for observations are derived from. Given the number of figures and material presented, I would add figure references to the end of each observation statement or conclusion to direct the reader to the supporting evidence. It was difficult sometimes to know where to look to corroborate the observation being made.</p>	<p>I think this is something of a stylistic question. Personally, I tend to find it annoying to read material in which every sentence ends with a (figure X). Particularly when an entire paragraph might refer to the same figure repeatedly. We have gone through the text with this comment in mind though, and tried to find specific places where it might be ambiguous. In those cases, we have added a few additional (Figure) references, as if it was difficult for our reviewers, it is likely to be difficult for many others too.</p>
Reviewer 3	Section 3.2.1	Second paragraph	Clerical	<p>Rows 2-4 of a figure are referenced without the figure number provided. Earlier, Figures 4 and 5 are referenced but it would be clearer and easier for the reader to be explicit as to which figure is being referenced. This is one example of many. I would revise using the guideline that any statement of fact or observation needs either a figure reference or a citation.</p>	<p>This is an interesting example. With the exception of one line that discusses Appendix figures, everything in this paragraph is about figures 4 and 5, and most of the references to rows and columns refer to both of them. It would feel awkward to me to put (figure 4,5) after almost every sentence, particularly since it would be in a parenthetical referring to rows and columns too. This just breaks up the mental flow for me too much. We have added a few references, particularly when a sentence referred specifically to temperature we reference figure 4, and when it referred to precipitation, we now reference figure 5. We hope this is sufficient to clear up the ambiguity, while still maintaining a smooth flow.</p>
Reviewer 3	Section 3.1.2	First paragraph	Clerical	<p>The statement is made that "both methods demonstrate similar patterns across the CONUS" but no figure reference is provided to support this statement.</p>	<p>We have added a reference to Figures 4 and 5 here.</p>

Reviewer 3	Page 43	last sentence	Clerical	the statement is made that “land surface wetness, soils, and vegetation clearly play an important role” but there is little evidence provided to quantitatively support this statement. I would soften this statement if it remains in the text.	We agree, this was overstated a bit given how difficult it is to see some of these features on a relatively small CONUS wide map. We have toned down this statement, and pointed out more specific examples that should make it easier to see a few examples. In addition, the next section on changes in extreme runoff also points to specific places where the land surface model is effecting the spatial patterns of change.
Reviewer 3	Section 3.1.2			The manuscript shifts between a variety of terms: LOCA and BCSD, LOCA-VIC and BSCD-VIC, and Maurer and Livneh. I would recommend more consistent and precise usage of these combinations. Consider Maurer-LOCA, Livneh-BCSD, Maurer-LOCA-VIC and Livneh-BCSD-VIC when discussing the combined uses of these products and datasets. An example of where this leads to confusion is in the second paragraph, last sentence of Section 3.1.2 2 where a comparison is made between LOCA and Livneh, which seems like to incorrect equivalence. Having this precise association explicit when discussing the products will also remind the reader about the differences in the forcing data, which is a main finding of this memo.	We agree that we have not always been as consistent with some of these terms as we should have. However, LOCA-VIC refers to something different than LOCA. LOCA-VIC refers to the VIC hydrology model simulations forced with LOCA meteorology. While Maurer and Livneh are distinct in that they are the observational datasets, so when saying Maurer-LOCA-VIC, I am not sure what that would be referring to. I'm also not sure what the problem is with comparing LOCA and Livneh. They are different datasets, even if one is trained on the other. In particular, we state "In this basin, LOCA is consistently drier than the Livneh observations across all GCMs, " and this would seem to be evident in the figure that is referred to. The LOCA-obs rows all show large brown (dry) regions in the upper portion of the Columbia river basin. We have added an explicit reference to figure 5 at the end of this sentence now.
Reviewer 3	Page 9	last sentence	Clerical	What is meant by “streamflow after runoff”?	We agree, this was not well worded. It now reads
Reviewer 3	Page 9	last sentence	Clerical	Change to read “we evaluate daily values...”	We agree that this was somewhat awkward, it referred to VIC grid cell runoff that had been routed through a river network to produce "streamflow". We have rephrased.

Reviewer 3	Section 3.3.3		Substantive	refers to “extreme changes” but it appears that only flooding is considered as an extreme, at least in the plots for this section. There is only a brief mention of “regional drought statistics”. The discussion of wet day fraction also seems to focus on extreme precipitation and not on lack of precipitation (which would be the drought side of extreme events). Therefore, I would either expand this section to look in more detail at drought or refine the section title and text to indicate the focus is on maximum streamflow and precipitation extremes.	In the interest of efficiency, we have retitled this to refer to extreme flood events only. Droughts would actually be quite interesting to look for the effect of the frequency dependent bias correction in LOCA.
Reviewer 3	Page 9	first bullet, 3rd s	Clerical	Would it more precise to use “produced” or “provided” instead of “saved”?	Save makes sense in the context that originally the modeling centers had more data, but they did not save that data (even internally); however, in a few cases, they may have saved more data, so provided is probably a better word. We have changed it.
Reviewer 3	Section 3.1.6		Substantive	I am not sure what this section adds. It interrupts the flow of the memo and I am not sure it helps in understanding the differences between the results. There would also need to be more support provided for many of the statements made in this section.	This was added initially at the suggestion of D. Pierce, and I think it adds a really valuable perspective. The bias correction differences have a large impact on the difference in the climate change signal for precipitation, and thus are important to at least discuss. This section has tried to note that there is no consensus in the community of which is "correct", but it does make reference to a few relevant recent studies that have investigated this, so I don't think it is unsubstantiated. I agree that the writing style interrupts the flow, so we have worked to modify the wording to better fit with the flow of the document, and repositioned it to follow the meteorology (as that is all that is really discussed in here).
Reviewer 3	Overall		Substantive	I wondered what the value of reporting the flow duration results are given they are not bias-corrected and the user is cautioned from using them. Would reporting the daily flow duration curves be representing results that are well beyond the predictive uncertainty? If this section is kept, I would write even stronger wording to caution users and enumerate more clearly the severe limitations of this data.	We have added some additional caveats. This section was added in part to illustrate just how much variation there is in these between BCSD and LOCA, while there is relatively little difference between GCMs, indicating that this is one statistic for which the downscaling method might have ended up being more important. There is also the connection to past reports that described similar basins.