Is the recent drought on the Colorado River "The new normal"?

A workshop on understanding the causes of the historical changes in flow of the Colorado River

September 24-25, 2018
NOAA Earth Systems Research Laboratory, Physical Sciences Division
Boulder, CO

Co- Sponsored by:

CIRES Weather and Climate Division
University of Colorado Boulder

Workshop report final version: 21 November, 2018

Disclaimer: The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the participants and authors of this report and do not necessarily reflect the views of NOAA or the Department of Commerce.

Workshop Attendees. Front L to R: Dennis Lettenmaier, Subhrendu Gangopadhyay, Connie Woodhouse, Julie Vano, Francesca Viterbo, Mimi Hughes, Rachel McCrary, Andrea Ray, Jeff Lucas, Dave Pierce, Klaus Wolter, Flavio Lehner. **Middle:** Chris Milly, Rob Cifelli. **Back:** Greg McCabe, Jim Prairie, Rebecca Smith, Xiao-Wei Quan, Jon Eischeid, Toby Ault, Joe Barsugli, Marty Hoerling, Robin Webb, Brad Udall. **Not in picture:** Balaji Rajagopalan, Ken Nowak, Ben Livneh, Tom Hamill. (Photo: Barb DeLuisi)

Authors of this report: Joseph Barsugli (CIRES, University of Colorado, Boulder) and Ben Livneh (CIRES and Dept. of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder)

Acknowledgements: Thanks to Christine Wiedinmyer of CIRES for co-sponsorship, and to Madeline Sturgill, Robbie Desen, Lucia Harrop, Chris Stuberg and Barb DeLuisi for workshop support.

Summary

Is the recent drought on the Colorado River "the new normal"? This question was the focus of an expert workshop hosted by the NOAA Earth System Research Lab Physical Sciences Division (PSD) on September 24-25, 2018 in Boulder, Colorado (See Appendix A: List of Attendees). The workshop was co-sponsored by the Weather and Climate Division of CIRES at the University of Colorado. More specifically, the workshop addressed the following: *What is our current understanding of the changing meteorology of the Upper Colorado River Basin (URRB) and the reasons for the hydrologic response to meteorological forcing?* The workshop was organized around four themes: The observational record; Temperature effects on runoff; Precipitation effects on runoff; and Other forcings. Presentations (limited to 4 slides total) were roughly grouped by theme, but discussion proceeded organically across all topics.

The timing of this workshop is opportune. Water Year 2018 naturalized flow at Lees Ferry will likely be in the bottom 10 percent of annual flows in the historic record, further stressing an already stressed system. The *Interim Guidelines for Lower Basin Shortages and the Coordinated Operations for Lake Powell and Lake Mead*, negotiated in 2007, are in effect until 2026. For the first time, stakeholders on the river are facing a real possibility that shortage conditions may be declared (in 2019), reducing deliveries to the Lower Basin by 500,000 acre-feet. Even now, Drought Contingency Planning (DCP) on the river is going on in order to plan for the consequences of a declared shortage on the river and the prospect that these shortages could get deeper in the period leading up to 2026. Furthermore, the Interim Guidelines specify that negotiations begin in the year 2020 on a successor agreement. Given the importance of management decisions on the River to the livelihoods of the seven Basin States, it is imperative that the current state of the science be clearly articulated as these negotiations begin.

Much of the discussion in the literature on potential reductions in UCRB streamflow has been framed in terms of sensitivity of runoff and streamflow to temperature and precipitation change -- the so-called "temperature sensitivity" and "precipitation elasticity." The results of a pre-workshop survey indicate that the participants came to the meeting with a wide range of estimates of temperature sensitivity, from a streamflow loss of 1.8 - 15 % per degree Celsius of warming. Likewise the pre-workshop estimates of "precipitation elasticity" — the percent change in flow for a percent change in precipitation — also were wide-ranging, from 1.2 - 3.0. Different methodologies used to estimate changes in streamflow generally fell in different parts

of the range of uncertainty. Analysis of GCMs -- whether from CMIP5 or in AMIP "Event Attribution" mode, tend to suggest temperature sensitivity at the lower end the range, offline hydrologic models tend to lie in the middle of this range, and regression analysis of observational data tend to lie at the upper end of this range. A key goal of the workshop was to document the current uncertainty in Colorado River flow sensitivity to meteorological forcing, and to motivate the community to better understand the reasons for these uncertainties by undertaking new observational studies and new model experiments.

Four sources of uncertainty were discussed at the meeting. The first three were discussed in detail: the "sampling uncertainty" due to the realization of only one sample of meteorology in the observed record, the uncertainty in the spatial details of temperature and precipitation in the basin, as evidenced by the difference in gridded datasets, and methodological uncertainty — for example the different treatments of soils among GCMs and hydrologic models, and the issue of coupling between land and atmosphere. A fourth source of uncertainty, the role of other forcings such as dust-on-snow or the changes in forest cover due to insect infestation, was touched on as well.

Breakout discussions on the second day were focused on developing consensus statements and identifying knowledge gaps. A summary of the main points from these breakouts indicates the following (more detail is provided in the main text and in Appendix C):

- The UCRB has warmed by about 1°C to 1.5 °C over the last century. The contributions of warming to streamflow decline remain poorly understood.
- Precipitation plays a larger role than temperature in drought, although its uncertainties are larger than temperature.
- Observational uncertainties, particularly between the early period (1900-1950) and the recent period (1985-2015) are the result of inconsistent sampling (fewer stations earlier, unequal sampling of high and low elevations).
- Observational uncertainties are especially large for determining precipitation change. These uncertainties make conclusive attribution statement on the role of temperature change vs. precipitation change on streamflow challenging.
- Models are necessary as part of the analysis, but open questions were raised into constraining them, e.g. with gridded observations, paleo data, or emergent metrics.
- The role of seasonality became apparent, e.g. cool-season precipitation is more important for flow than warm season
- The question of detectability of a true change in hydrology of the UCRB was raised : is a century of historical data are adequate to distinguish variability from underlying change?

- The issues of spatial scale is still not completely resolved. What spatial scale (e.g. horizontal resolution) is adequate to realistically model the basin and are GCMs and regional climate models reaching these resolutions?
- Despite the importance of temperature and precipitation, other changes have been occurring on the basin that could affect streamflow that are rarely considered by major analyses: dust-on-snow, tree mortality, surface water/groundwater interaction.

The pre-workshop survey sought discussion questions for the workshop for each of the workshop discussion themes. These are summarized in the report and in Appendix B -- and are worth noting here for their deep insights as well as breadth of scope. These discussion questions, reframed as research questions, could help define a research agenda for the UCRB. Being only 1.75 days long, this workshop raised more questions than it answered, and focused our attention on the divergent nature of our research results and the need to explain these differences. Yet there was also a sense of both urgency and excitement that many new strands of research could contribute to a better overall understanding of where the basin has been and why, to provide a sound basis for projecting its future. At the workshop, and in the post-workshop survey, there was near unanimous support for the idea of seeking a broader-based AGU Chapman Conference on the science of flow in the Upper Colorado River Basin in about two years' time to help with synthesizing the science and communicating it to stakeholders.

Purpose and Scope

The purpose of the workshop was to document the state of knowledge on historical changes in flow due to long-term climate change and internal variability in the Upper Colorado River Basin (UCRB, at and above Lees Ferry), including the changing nature of droughts. The question is framed around an inquiry from Eric Kuhn (of the Colorado River Conservation District) to Marty Hoerling about the low flows seen on the river since the year 2000: “Is this a drought or is it aridification?” In other words, are the conditions of the past 18 years “The new normal”?

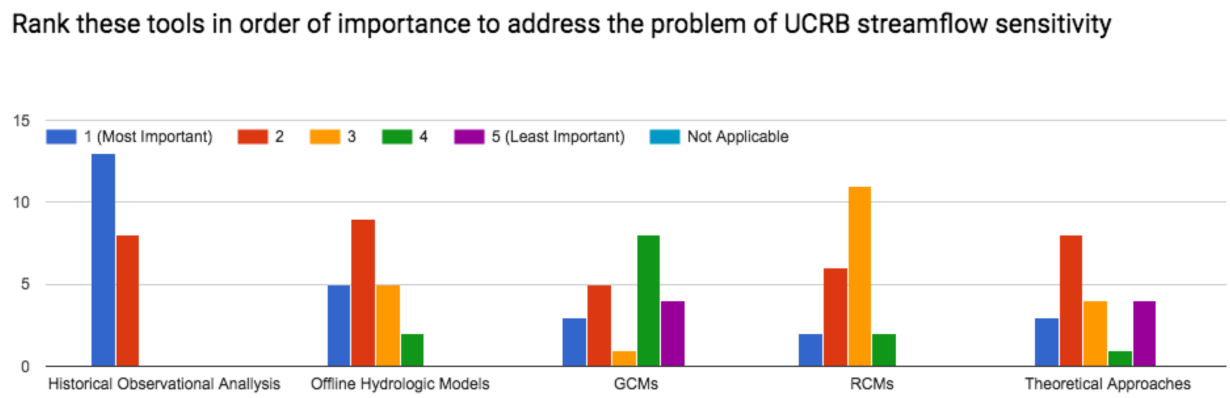
The scope of the workshop was limited. The workshop focused on the science of detection and attribution of historical hydroclimatological changes in the Upper Colorado River Basin (UCRB), explaining physical reasons why these changes happened, and comparing different methodologies. The aim is to better integrate information from observational analyses, hydrological and land-surface modeling and climate modeling. Given the short duration of the workshop (1.75 days) the focus was NOT on paleoclimate or on projections -- those could be, and have been, topics of workshops in and of themselves. Nor were stakeholders a large part of the conversation -- intentionally, so that the science could be presented without concern for communicating consistent narratives, but rather for exposing differences and generating research questions to reconcile or explain these differences.

Results of Pre- Workshop Survey

We asked workshop attendees to answer an online survey before the workshop to help stimulate and guide the discussion during the workshop (Appendix B: Pre-workshop Survey). A total of 23 unique responses were received out of the 26 invitees polled.

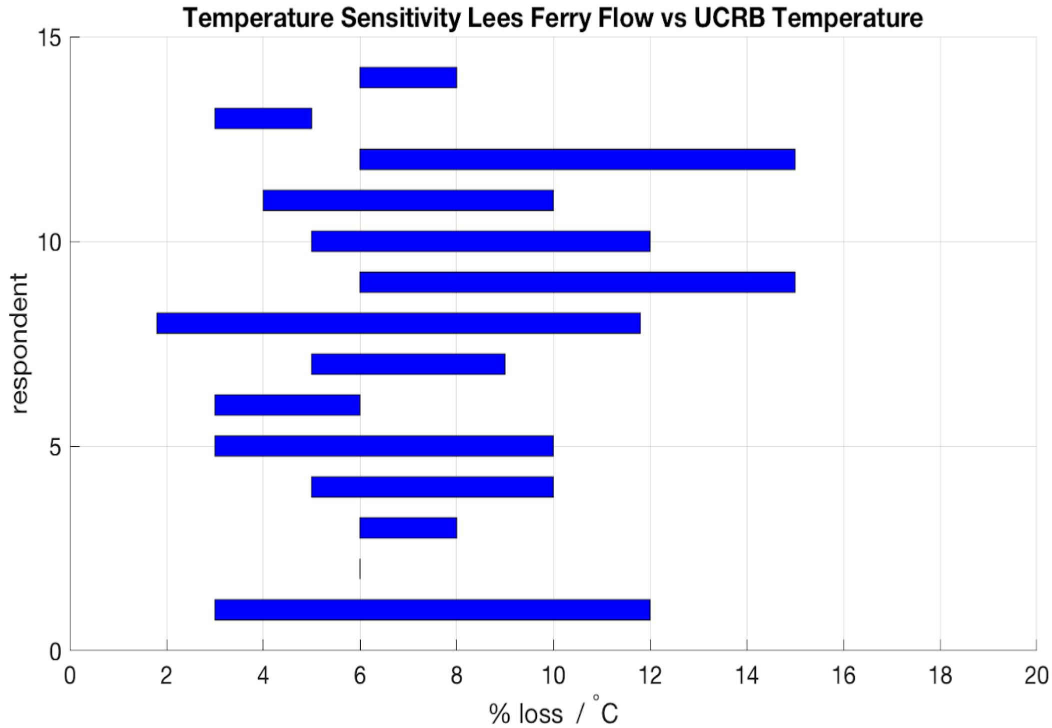
Physical Climate and Hydrology Questions

Because one of the goals of this workshop was to bring together research using different methodologies, the first science question addressed this topic:

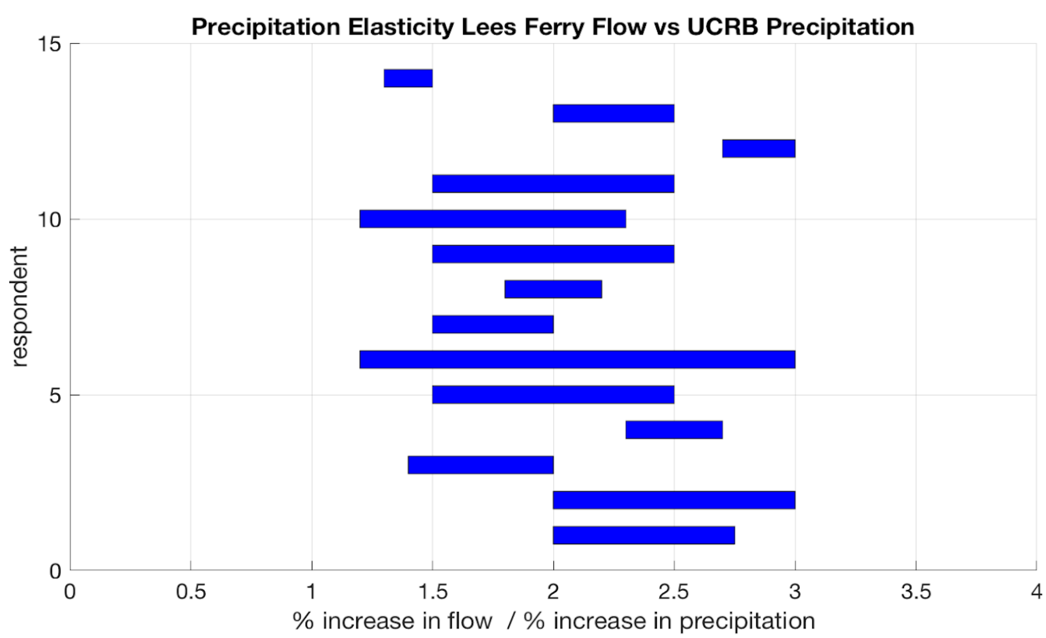


Clearly, historical observational analysis was seen as fundamental to understanding the basin. Offline hydrologic models (meaning that they are not coupled to an atmosphere or general circulation model) have been the workhorse models for understanding and projecting flows in the basin, and these were also popular. The importance of GCMs was, interestingly, split into two camps among the attendees. RCMs fared better than GCMs, presumably because of their greater resolution in the basin, even though RCMs have not been widely analyzed for this purpose. Theoretical approaches (such as Budyko methods, very idealized models, and sensitivity analysis) was more widely spread in its importance, perhaps reflecting the fact that it was not well-defined in the questionnaire. Given the option to offer other tools and analyses the respondents listed the following: specific analysis of high-runoff producing elevations in the basin, the emerging modeling tools of variable resolution GCMs and the National Water Model, as well as paleoclimate/paleohydrology analyses and robust/flexible decision-making alternatives, which were not the focus of this workshop.

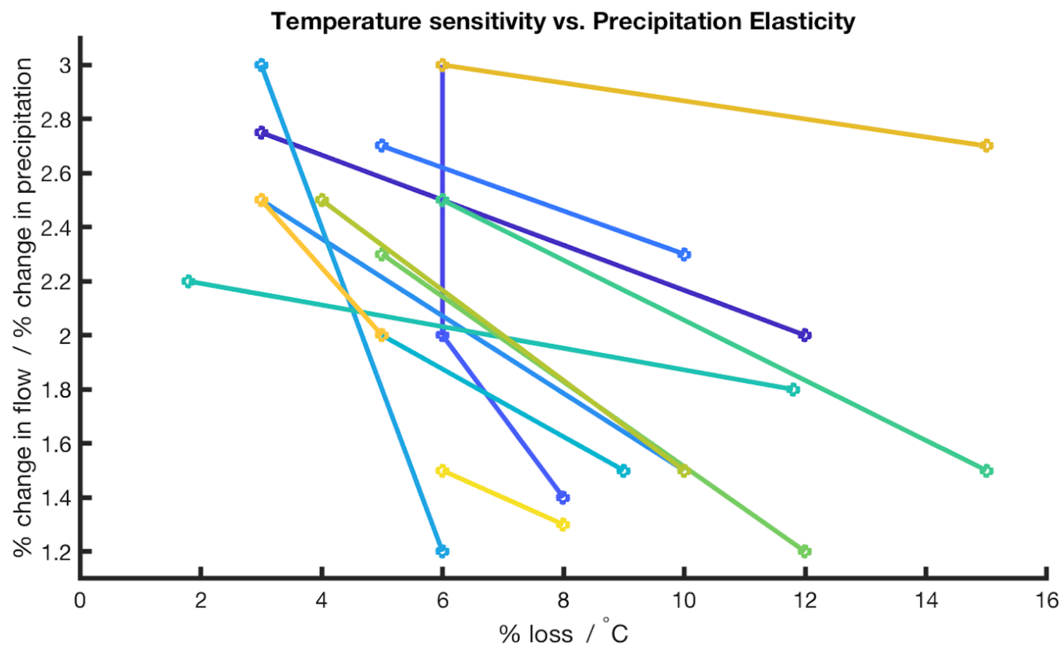
The next question surveyed the professional opinion about the range of sensitivity of the overall natural flow in the UCRB at Lees Ferry to long term temperature increase. The responses for each individual are shown as a horizontal bar from the low-end to the high-end of their expressed range:



The same question was asked for precipitation elasticity (the %-change in runoff for a %-change in precipitation in the whole basin):



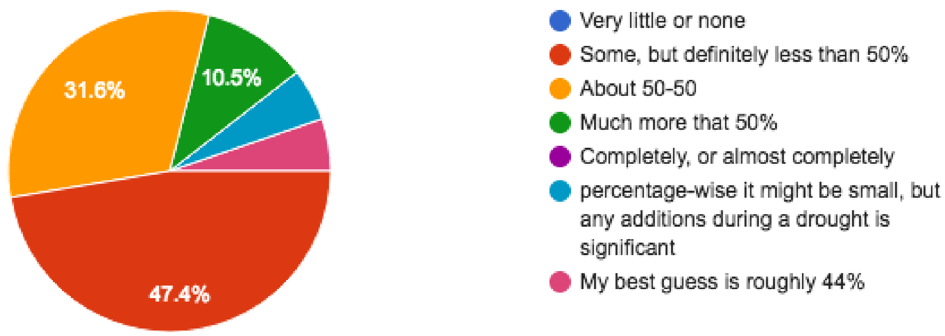
The expert opinions of the group range from low values of about 1.2 to 3.0, a remarkably wide range. A very interesting feature emerges when the ranges for the two sensitivities are plotted against one another. Each line joins the minimum and maximum values for a single respondent:



It is generally recognized that there is a tradeoff between precipitation and temperature sensitivities, and this is indicated by the slopes of the lines. It is remarkable that a large fraction of the group sees approximately the same tradeoff between these two quantities, even though there is a large range in uncertainty in each quantity individually.

The final question addresses "the new normal" directly.

In your view, what portion of the recent drought (say, 2000-present anomalous streamflow) is an anthropogenic signal of change on the river?
19 responses



While this question was not intended to be authoritative it is noteworthy that no-one answered "Very Little or None". 42% answered about 50% or much more.

Workshop Discussion Questions

Many potential discussion questions were suggested in the Pre-workshop survey. A short summary is presented here; the full set is in Appendix B.

Observations: There were many questions about uncertainty in our observed datasets, including naturalized flows, and gridded precipitation and temperature. The potential role (and dearth of observations) for groundwater was also raised. The need to focus specifically on higher elevations of the basin, including the role of snow processes was emphasized, as well as the observational challenge this represents. Reconciling observational estimates of temperature sensitivity with models, and in general using observations to shed light on models was suggested, as well as using the paleo record to put the historical record in context.

Temperature Effects: The interpretation and potential usefulness of basin-wide temperature sensitivity was questioned. The roles of seasonal, spatial, and temporal variability were emphasized. Are we fully taking into account co-varying temperature and precipitation? The need to better understand why different methodological approaches give different answers was mentioned. Several questions centered on the energy balance in the basin as a key to understanding the water balance, including snowmelt and runoff efficiency.

Precipitation Effects: As with temperature, the spatial, seasonal, and temporal variability were cited as important, as well as multi-decadal variability. The role of snow processes (vs rain) in precipitation efficiency in the basin was also raised. Many of the questions focused on understanding the precipitation trend itself, not just the sensitivity, including how to reconcile GCM-modeled precipitation trends trends with observations. Downscaling issues came up particularly for precipitation (rather than temperature).

Other forcings: Other forcing factors including CO_2 fertilization and water use efficiency, vegetation change (including beetle-kill), dust radiative forcing: Have these had an impact already on the historical records that should be taken into account? Many projection methods (including offline hydrologic models) assume static vegetation and other forcings but is this realistic in the future?

Other questions: There were several questions that suggested a deeper look at our methodologies, including differences among land surface model formulations in offline hydrologic models and in GCMs. Can we say anything about the quality/accuracy of the CMIP model simulations? And finally, how can we make the science align more with management needs.

Workshop Discussion Sessions

Scheduling of the slides was flexible throughout the two days, but was grouped, more or less, by the main workshop themes. To encourage discussion and reduce “powerpoint fatigue” participants were limited to four powerpoint slides, total, for their presentations. The proposed discussion themes were as follows:

Monday AM	Theme 1: What are the primary uncertainties in the observational record for UCRB precip, snowpack, temperature and streamflow? What are the implications for detection of historic hydroclimate changes.
Monday PM	Theme 2: Temperature-streamflow physics: How is the magnitude of the temperature “effect” on streamflow estimated, what are the uncertainties, and (how) can these be reduced? How well do observations constrain model estimates?
Monday PM	Theme 3: Precipitation-streamflow physics: How is the precipitation elasticity of streamflow estimated, what are the uncertainties, and (how) can these be reduced? How well do observations constrain model estimates?
Tuesday AM	Theme 4: Other forcings--bit-players or lead actors?: greening; dust; insect; fire, other?: do these influence our view of the past 120 years of change in flows? [Ed. This Theme was only touched on in the workshop]
Tuesday PM	Small Group Breakout: What consensus has the group reached? What are major research gaps?

A summary of the presentation topics organized by presenter, not necessarily in the order presented, follows.

Monday Discussion

After a welcome by Robin Webb, Marty Hoerling opened the meeting by recalling the questions that were raised at the end of the 2014 BAMS article on Understanding Uncertainties in Future Colorado River Streamflow (Vano et al, 2014), citing the major sources of uncertainty in streamflow projections: 1) Global Climate Model (GCM) and emission scenario selection, 2) Spatial scale and topographic dependence of climate change projections - simulations of land *and* atmospheric processes, 3) Land surface representations, and 4) Statistical downscaling methods.   All of these sources of uncertainty (with the exception of scenario selection) have their counterpart in understanding the historic flows as well.   Marty also presented work on an "event attribution" approach to understanding the sensitivity of the river to temperature and precipitation long-term changes.

Connie Woodhouse discussed analysis of the paleo record, finding at least six 19-year periods since 1400 that exceeded the severity (in terms of cumulative deficit) of the recent drought on the river. She also discussed paleo reconstruction of runoff efficiency, finding evidence for lower runoff efficiency during many past droughts. Paleo reconstructions of runoff efficiency can shed light on the role of temperature on streamflow during past warm and cool droughts.

Greg McCabe related the trend analysis of precipitation, temperature and streamflow from McCabe et al. (2017) showing the observed higher-than-expected streamflow pre-1930 and lower-than expected streamflow post-1990 (for 10-year running means).

"Klaus Wolter discussed precipitation-flow relationships for the full SNOTEL record from 1979 to 2018. He found that near-normal precipitation over the SNOTEL domain produced 10% less runoff in the recent half of the record (1999-2018) compared to the early half (1979-1998). At least some of this discrepancy can be explained by a preference of La Niña over El Niño events in recent years, since the latter tend to produce more runoff for the same precipitation anomalies than the former. On the other hand, he also raised the issue of confounding effects from the pine beetle epidemic that has left a legacy of standing dead trees without needles ("gray stage") which may have generated up to 5% more runoff than otherwise expected."

hris Milly presented some work on estimation of mean basin precipitation, WITH CONSIDERATION OF measurement error and spatial sampling error. Applied to the UCRB it shows a substantial ESTIMATION ERROR compared to the annual standard deviation. He also discussed his recent paper on the sensitivity of streamflow to temperature using a Budyko-type framework, noting that, globally, most temperature sensitivity values are low (on the order of -1 to -3 %/°C).

Joe Barsugli presented slides between running the show, and discussed uncertainty in the naturalized flow record, sampling uncertainty in regression-based estimates of temperature sensitivity, and also using the energy balance as a unifying metric to understand differences in GCM (coupled energy balance) and off-line hydrologic models (usually a specified energy balance from empirical estimates such as MTCLIM).

Balaji Rajagopalan discussed a statistical model of variability, motivated by wavelet analysis of paleo-reconstructions, that could potentially give insight into whether we are likely to see increased or suppressed variability on the river.

In a related vein, Subhrendu Gangopadhyay discussed some work on predicting regime shifts in flow on the river based on paleo streamflow reconstructions.

Afternoon discussions followed:

Jim Prairie gave us an update on the process for the Interim Guidelines, and the ongoing Drought Contingency Planning (DCP) efforts. He presented two alternate futures, one derived from all available past hydrologies (1906-2015), and a "stress test" hydrology considering only 1988-2015 traces. The "stress test" hydrology poses a much greater risk of Lake Mead falling below 1025' elevation *before* the Interim Guidelines expire in 2026. He also gave a summary of Reclamation-sponsored research in the basin to develop a better knowledge base for upcoming decisions. [Ed. note -- According to the discussions at the meeting, this presentation clarified the ongoing process of DCP and the upcoming re-negotiation of the Interim Guidelines for researchers. It also emphasized the need for short-term risk assessment (a couple of decades), potentially including rationales for different "stress test" hydrologies, in addition to long-term projections. The timeliness of our research was clear, and also the limited time frame in which we will have to get it done to be included in the process of re-negotiation.]

Flavio Lehner discussed his recent work in Earth System Models, and in particular looking at CMIP5 models, and the NCAR CESM Large Ensemble. Estimates of temperature sensitivity from most CMIP5 models were quite low, with some models even exhibiting the opposite sign of sensitivity (flows increasing with warmer temperature). "Since biases in model sensitivity and elasticity appear to directly affect future runoff projections, understanding and alleviating these biases is an important step to increase confidence in future regional hydrologic projections from ESMs."

Dave Pierce showed his work on evaluation of CMIP5 models, focusing on the question: “Are CMIP5 models, as a whole, wetter than CMIP3 for the UCRB?” Using a skill-based weighting, the raw CMIP5 models do indeed project a wetter future than CMIP3 models, and the weighted mean has a small increase in precipitation by the end of the century (RCP85). He also exposed a particular assumption in the VIC formulation of evapotranspiration (about nonlinear temperature dependence of canopy resistance) that may account for some non-intuitive results -- more as a cautionary tale that the details of models can be important.

Tuesday Discussion

Brad Udall focused on sequences of low flow years of length from 2 to 18 years, looking at 2000-2017 compared to the 20th century. The 21st century low flows sequences are as bad or worse than any in the 20th century. He was also concerned with the ability of the CMIP5 GCMs to capture the aridity of the western US, as indicated by a simple index, the Aridity Index (P/PET, from Seager et al, 2018), and hence was concerned that CMIP5 was a dominant source of regional projections.

Dennis Lettenmaier summarized the key result from a recent paper (Xiao, Udall, and Lettenmaier, 2018) whereby the trends over the last century are decomposed into a precipitation-forced and a temperature forced signal, concluding that a little over half of the annual runoff trend was caused by the temperature trend. According to this analysis, warming also contributed over 50% of the streamflow deficit during the 2000-2014 “Millennium” drought.

Rachel McCrary showed some preliminary results from analyzing CORDEX-North America regional climate models at 25 km resolution. Indications are that these models share the tendence for the forcing GCMs to have a large wet bias in the basin. Further analysis is needed.

Mimi Hughes discussed the prospects of using a weather-scale model such as WRF at very high resolution (a few km) to estimate precipitation (and snow) in mountainous areas -- perhaps better than interpolation from station observations. This is an intriguing possibility for the Colorado River basin, particularly for analyzing droughts and pluvials in the relatively data-sparse era before 1980 (the onset of SNOTEL measurements).

Ben Livneh discussed some issues that may come up in the basin regarding "other forcings" in the basin beyond temperature and precipitation. The prospect that dust-on-snow may have influenced the water balance was raised, though the magnitude of *trends* in this effect were questioned. Land Use Land Cover (LULC) changes are modeled quite differently in different datasets, and this may have implications for interpretation of the CMIP5 and CMIP6 model simulations. Issues of water use efficiency and other biophysical changes may also play a role. We could design experiments using a modeling system such as NCAR CESM to get at the order of magnitude of these effects.

Julie Vano finished the Tuesday presentations with a reprise of the 2014 paper "Understanding Uncertainties in Future Colorado River Streamflow". This paper documented the state of the science in the early 2010's. The talk provided the context for the breakout sessions in the afternoon, and a baseline to measure whether our consensus on the changes in the river had changed since then. It was also notable that while that paper focused primarily on projections, many of the questions were also germane to understanding the historical changes. It is worth repeating the conclusions from that paper here:

- Temperatures (T) will rise over the coming decades
- Precipitation (P) less certain, but will likely decline annually
- T and P responses depend on intensity of greenhouse gas emissions
- Warmer T (ignoring P) will reduce runoff production (our estimates -6.5 ±3.5% per °C at Lees Ferry)
- Change in P results in streamflow response of 2 to 3 times (5% decline in P results in 10-15% decline in streamflow)
- Coarse spatial resolution does not resolve high elevation hydrologic processes that dominate basin runoff production
- Natural variability in paleoclimate reconstructions show prolonged multi-decadal dry periods. Going to see decades of sustained streamflow much lower than observed in ~100 years instrumental record (especially with climate change exacerbation)

Summary of Breakout Sessions

We held small-group breakout sessions in which the participants were asked to respond to the following questions: *What consensus statements could they make?* and *What are the knowledge gaps and research strategies to fill these?* Below is a summary of general topics in terms of key drivers of drought and change, their estimation (observations versus models), and their uncertainties.

The role of temperature

There was broad consensus that temperature in the UCRB has warmed in the last century, with warming of greater than or equal to 1C. The warming trend is nonlinear (i.e. warming from 1900 – 1940, flat from 1940 -late 1970s, warming to present). It was generally accepted that temperature has contributed to low flow during the Millennium drought, primarily by reducing the runoff efficiency. The magnitude of this effect, the importance of non-linear warming, as well as the relative of the uncertainties were not conclusive.

The role of precipitation

Precipitation was generally deemed the largest driver of Colorado River flow changes over the last 100 yrs. There was a sense that the seasonality of precipitation is of importance and that cold-season precipitation is main driver of flow, which is something that has come out of hydrologic model sensitivity analyses, as well as correlation analyses between winter precipitation and water year flow. Some rationale for this assertion was discussed, for example, during cold-season months:

- There is less ET than during summer precipitation due to colder temperatures and less radiation
- Soil moisture is higher, or frozen so there is a higher runoff efficiency
- Residence time in soils is shorter during rapid snowmelt, which supports runoff

Despite the importance of precipitation, neither the magnitude nor sign of precipitation change in the UCRB over the last century are well known. The magnitude of the true change is likely within +/-20%

Observational data and sampling issues

There was consensus that the annual streamflow at Lees Ferry has declined over the last century by as much as 10 to 15 %. The "Millennium Drought" post-2000 is seen as the most sustained drought in terms of consecutive (cumulative) low flows during the instrumental period [Note the term "cumulative" based on discussion at the time]. Paleo data (1400-present) reveal several sustained droughts of equal or greater severity than the Millennium Drought [see C. Woodhouse slides for precise definitions]

Observational data represent a major limitation to our understanding of changing conditions. For example, the selection of different observational stations by commonly used datasets can affect the inferred meteorological trends, more important/divergent for precipitation than temperature. Further, observational data are seen as less reliable prior to key time periods, e.g. 1948 (digitization), 1979 (beginning of SNOTEL). However, temperature anomalies are better represented than precipitation because they vary across larger scales, but CO-OP stations may have declined in mountainous areas, so these can't be assumed to be monotonically increasing. Lastly, spatial differences inside the UCRB are extremely important – averages over the basin need to be supplemented with a finer spatial scale examination. Overall, any trends (flow, precipitation, temperature) depend on the endpoints and similarly the basin sensitivity depends on timescale.

The role of models

An attempt was made to reconcile the use of models to address the problem. It was felt that while GCM model ensemble averages generally get the warming trend right (within several tenths of a degree C), uncertainties in temperature and precipitation feedbacks together with differing model parameterizations limit our ability to disentangle how much T and P have forced historical streamflow changes. In using LSMs to hydrologically 'downscale' GCMs, temperature sensitivities become very broad. Over the UCRB, CMIP5 models are wetter than CMIP3, but large biases (typically wet biases) relative to observed hydrology remain.

Research Gaps

A helpful question was raised to frame the problem: is the Millennium Drought truly a drought, or the beginning of aridification? To address this question a set of research gaps were identified that include understanding how much things have changed relative to earlier periods, why things have changed, on the use of models and observations (e.g. the degree to which the basin, and the roles of seasonality and other non-atmospheric drivers, e.g. dust, tree mortality, and groundwater.

First, putting the current drought in context and understanding *how much* things have changed an open question was raised as to how wet the early period (1900-1925) was relative to the current period (e.g. McCabe et al., 2017s). The workshop revealed that key uncertainties in observational datasets, primarily in precipitation and natural flow, limit our ability to adequately answer this question. Some ideas to address this gap included a more rigorous uncertainty analysis of gridded datasets, revisiting early mining precipitation records, using tree ring data, dynamical or statistical downscaling of 20CR/ECR re-analyses, as well as applying Bayesian inferred precipitation from streamflow.

Next, many research gaps were identified in terms of *why* things have changed. This mainly centered on gaps in our current understanding of basin sensitivity. For example, there is a notable difference between observationally and model based estimates of UCRB streamflow sensitivity to temperature and precipitation. Model consistency was a key issue that was raised as well. A few ideas to address the issue were to (i) assess inter-model consistency, in terms of variance, (ii) to use the historical record to constrain model application, (iii) to force all the LSMs from the CMIP6 ensemble with observations and evaluate their performance over the historical period.

A suggestion was made for an assessment of potential predictability of the annual flow on the UCRB on different timescales. It would be useful to define "emergent" metrics (i.e. metrics outside those to which the models are 'tuned') for evaluation and constraining ESMs GCMs, LSMs etc. Sensitivity and elasticity are examples of emergent metrics. The predictability assessment could include AMIP-style simulations, initialization of forecast models (decadal prediction studies), and statistical/dynamical systems analyses

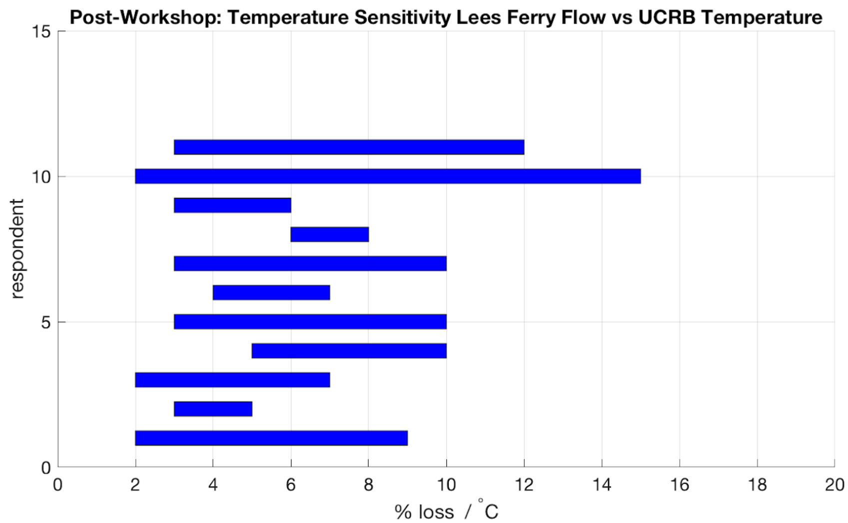
On the observational side, gaps were identified into how elasticity, runoff efficiency, and sensitivity vary by sub-basin, as well as a larger question about which spatial resolution do we really need to model the UCRB? An idea was raised to try to get paleo estimates of elasticity, efficiency, and sensitivity, in order to put the observational and model estimates into context. To address the role of seasonality, a suggestion was put forth to use our 'best observed' time period (1980-2010) as a control and contrast other time periods against this time period, with both observations, and models.

Lastly, the notion that perhaps we are being misled by attributing all of the non-precipitation driven changes in the UCRB to temperature, in light of other changes that have been occurring: dust-on-snow, tree mortality, surface water/groundwater interaction, and additional unknown forcings?

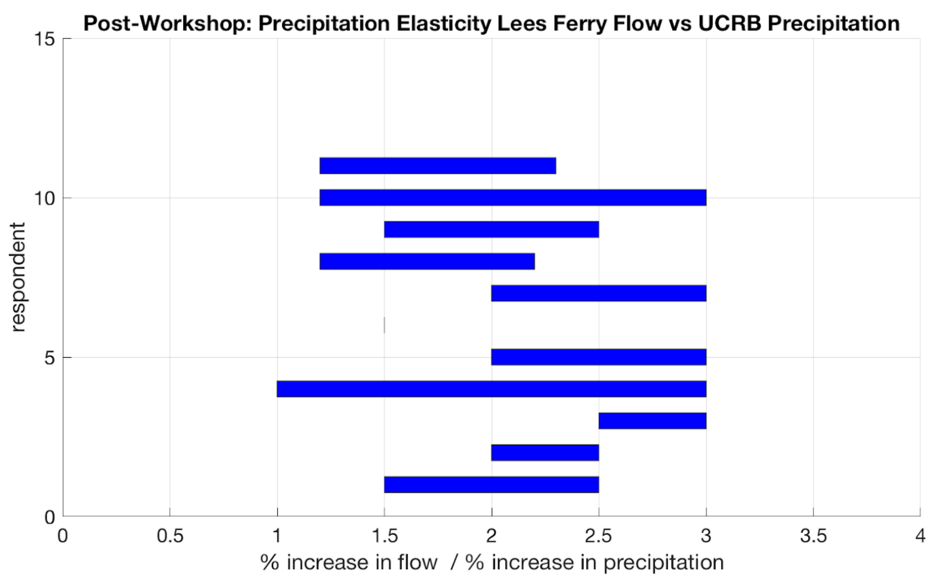
Results of Post-Workshop Survey

The post workshop survey had 12 respondents, about half the number for the pre-workshop survey. We repeated the temperature sensitivity, precipitation elasticity, and "anthropogenic signal" questions.

The graph of all respondents for temperature sensitivity is shown here:

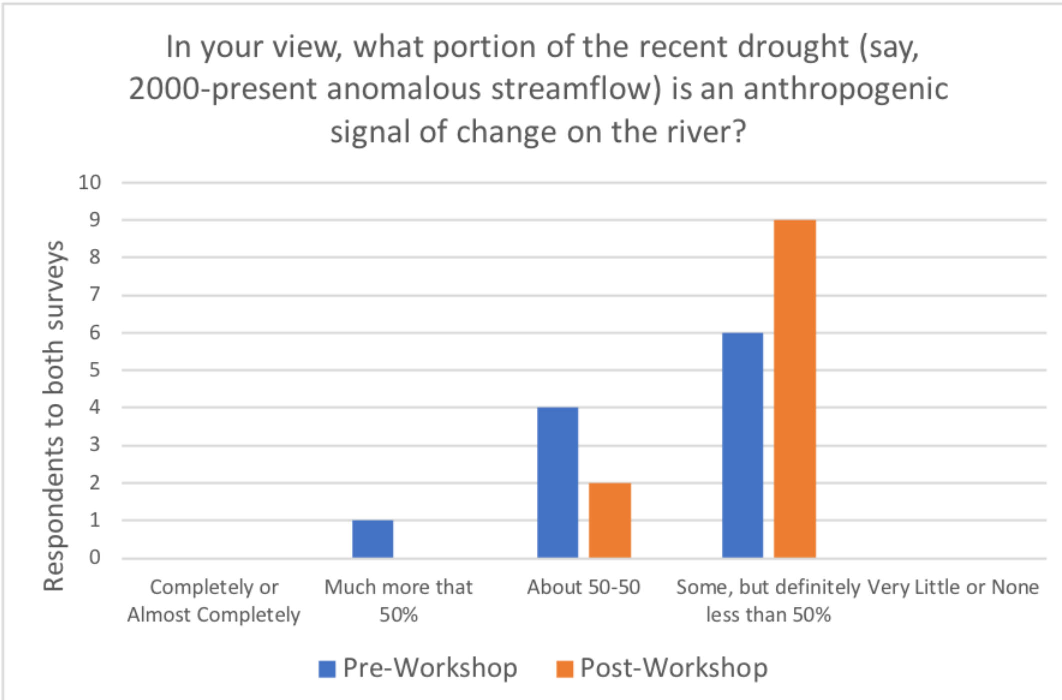


Of the seven respondents in common for temperature sensitivity between the pre- and post-workshop sessions, three expanded the range of uncertainty on both ends, while one expanded the range only on the lower end of sensitivity.   For precipitation elasticity we see the following:



Of the seven common respondents for the precipitation elasticity question, two expanded the range in both directions, one expanded the range on the upper end, one reduced the range on the lower end, and three stayed the same.

For the question of anthropogenic signal in the recent drought, respondents shifted towards "Some, but definitely not 50%" from "About 50-50, compared to the Pre-Workshop survey.

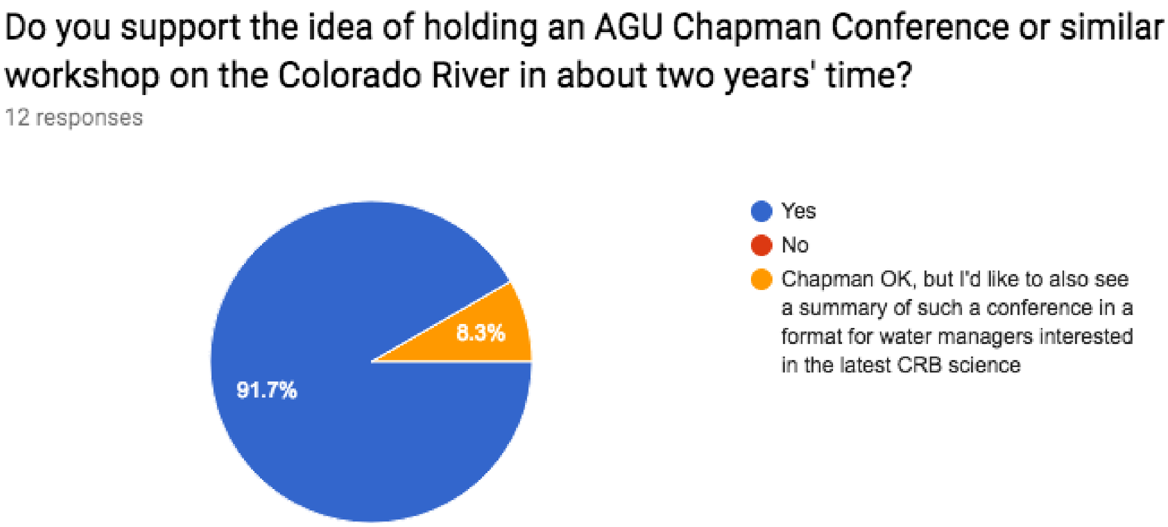


Respondents indicated how their view of the overall science had changed, mostly indicating that there was not much change (and that the initial large range of sensitivity had not been narrowed):

- Temperature sensitivity went down based on results presented and discussion.
- no meaningful change
- I don't think I've changed, but can't fully remember my previous submission.
- I deliberately did not look at my old responses, but I don't think they are very different.
- a little, range expanded based on seeing a variety of approaches (though confidence that range in ballpark also increased)
- I think I've widened my estimates a little bit. I wasn't aware of just how large the uncertainties in the observational record of temperature and precipitation are.

Follow up to this workshop

The organizers broached the idea of holding a future science-focused workshop, such as an AGU Chapman Conference, in about two years' time as a focal point to devote more time to these questions. This would bring together not only the work on attribution of the historical record, but also the paleoclimate and projections, as well as stakeholder perspectives. A model for this might be the Chapman Conference on the California Drought for which Marty Hoerling was one of the conveners. There was overwhelming support for doing this expressed both at the meeting and in the post-workshop survey. The Proposal for Chapman Conferences are due in 15 March 2019 and the conference must be held a minimum of 15 months later.



Other ideas for keeping the momentum rolling on this included:

- A manuscript summarizing the outcomes would be interesting. Probably many 'experts' would be willing to contribute.
- EOS article, form a program committee that enlists folks from workshop, draft a proposal of scope and overarching questions

- more frequent/incremental seminars that focus on methods and assumptions before there are results (almost impossible for practitioners to follow along when everything is crammed into 50 minutes)
- It seems like the workshop didn't converge on any sort deliverable, which could have provided a focus for ongoing collaborations
- We need something; might be as simple as an email list. AGU Session?
- EOS article about this workshop
- funding the could support updates to reconciling different approaches, look more across space and time (not fixed on Lees Ferry in past 30 years)
- Maybe a Nature Climate Change perspective summarizing the workshop to make the problem widely known and stir interest in a potential Chapman conference

We, the workshop organizers, are preparing a short workshop summary for EOS (Note: these are limited to two or three authors).

Workshop Feedback

The respondents were unanimous expressing their enjoyment of the workshop, and the "open" structure and format led to it being a "fun" 1.75 day meeting. The insights from the USBR on how this information may fit into the revision of the Interim Guidelines was singled out as a useful guide, though one person thought that it might have been possible to "articulate some overarching questions to provide a bit more focus, which also may have resulted in some material for a synthesis or perspectives paper." We hope that as we, as a group, remain involved in the critical questions that face the basin, that such a synthesis could emerge.

- Thank you for such a nice job!
- Insights from USBR on the need for science input to the revised Interim Guidelines process very helpful
- really enjoyed it, thanks for organizing
- The workshop was quite interesting and it was good to hear the different results and perspectives. The structure was open, as appropriate for this type of workshop, but it might have been possible to articulate some overarching questions to provide a bit more focus, which also may have resulted in some material for a synthesis or perspectives paper.
- Thanks for pulling this together; it was valuable.
- Great workshop, thanks!

- I liked the format after all, and looking forward to summary of the end of workshop poster boards.
- Thank you. It was a fun meeting.
- It was among the best 1.75-day long workshops I've been too - thanks so much! [Ed: that depends on how many 1.75 day workshops that person has been to!]

Appendix A: List of attendees

Name	Affiliation
Connie Woodhouse	Univ. of Arizona
Dennis Lettenmaier	UCLA
Dave Pierce	Scripps Institution of Oceanography
Jim Prairie	USBR
Brad Udall	Colorado State Univ.
Julia Vano	NCAR
Greg McCabe	USGS
Flavio Lehner	NCAR
Balaji Rajagopalan	Univ. of Colorado
Chris Milly	USGS
Jeff Lukas	Univ. of Colorado/WWA
Toby Ault	Cornell Univ.
Rebecca Smith	USBR
Rachel McCrary	NCAR
Ben Livneh	Univ. of Colorado
Marty Hoerling	NOAA/PSD
Joe Barsugli	Univ. of Colorado/PSD
Klaus Wolter	Univ. of Colorado/PSD
Xiao-Wei Quan	Univ. of Colorado/PSD
Mimi Hughes	Univ. of Colorado/PSD
Rob Cifelli	NOAA/PSD
Tom Hamill	NOAA/PSD
Robin Webb	NOAA/PSD
Francesca Viterbo	Univ. of Colorado/PSD
Ken Nowak	USBR

Appendix B: Pre-Workshop Survey Questions for Discussion

Observations

Flow

- How reliable are the **naturalized flow** reconstructions?
- Empirical estimates of sensitivity usually assume negligible **groundwater storage** change (so P=ET+runoff). Is this an acceptable assumption?
- Are observation based methods for estimating T sensitivity missing other portions of the hydrologic cycle - e.g. changes in **GW** contributions to flow?

Snow/Elevation

- Are a dearth of high elevation observations limiting our ability to answer the questions at hand?
- What approaches to better obs might be pursued? Are we leaving some things out? How much does our ignorance of higher elevations matter?
- How much value would be added by expanding Snow Observations beyond the current SNOTEL network?

Basic uncertainty/gridding

- Gridded data sets are really important for modeling, yet they are all over the map. Do we need to look at these datasets in more detail?
- Uncertainties in precipitation for early record (less so for temperature/runoff)
- How good are we at estimating UCRB P and T from historical point observations? How much do errors in estimates of historical UCRB P and T contribute to uncertainty/error in sensitivities?
- What are the effects of low frequency (e.g. decadal to multi-decadal) variability on future UCRB flow? and are these effects as large or larger than the effects of global warming?

Determining Elasticity, etc.

- Are the model estimates of elasticities consistent with the observations? When I look at this, I think that the observations show lower precipitation elasticity than most models, or at least on the lower end of the model range, something like 1.8-2.0. If the USBR is counting on wetter conditions to overcome the temperature-driven flow deficits, this is an even bigger stretch than we thought. And it already seemed implausible to begin with.
- How well can we estimate P elast. and T sens. from observations alone? How long of a record would we need? Can large ensembles with climate models help us estimate how

far off we are from being able to constrain things with the instrumental record?

- What are the minimal observational requirements for meteorological measurements of P and T (#of stations, station locations, historical duration, accuracy of measurements) to determine CR runoff sensitivity to met forcing?

Paleo

- Modern drought in paleo context. What aspects conspired to produce paleo droughts?

Temperature Effects

Basics

- To what extent is temperature a driver on a large basin?
- Is temperature sensitivity even a useful concept?
- Is temperature sensitivity of runoff an inherent property of a river basin, and of so, how much (and which) data are needed to determine the temperature effect?

Variations

- How does temperature effect vary spatially? temporally?
- What is the elevation dependency of temperature effects on runoff?
- Hoerling et al. present a different result regarding the role/importance of temperature and precipitation on flow reductions relative to other results; it would be useful to explore the sources of the differences and the data uncertainties related to the different approaches.

Covarying T and P

- How do T&P covary? We have an artificial distinction between these two, mostly to make it easy to analyze. But they are tied together -- higher precip will push T effects down, and lower precip will make them worse.
- How much of recent warming is due to recent dryness?

Energy Balance

- How accurate at the LSM energy balances?
- Can we constrain T sens estimates with an energetics perspective? E.g., how much more water can you actually evaporate with 1C of warming? Is this in the ballpark of what the flow declines amount to?
- How well do we understand the surface energy balance of the UCRB? Does it matter?
- Role of warming in snow-melt efficiency
- What is the role of snow albedo feedback in amplifying temperature sensitivity? (Less snow-->more radiation-->more ET--> less runoff?)

Precipitation Effects

Basics

- What about the high-res simulations where we have increase in precip efficiency?
- How does the progressive change of pcpn from snow to rain (e.g. due to warming) affect the runoff ratio in the UCRB?

Variations

- How much does spatial pattern of precip matter?
- How do precipitation effects vary spatially? Temporally?
- Is the P elasticity of runoff different on annual versus centennial time scales?
- Multi-decadal variability of precipitation and the role of anthropogenic forcing
- What is so special about the early season (OND) precipitation - why does it 'punch above its weight'?

Interpretation of GCMs/methodologies

- CMIP5 MMM says no forced precip trend. Some models in there might be crap. Can we do better than MMM on this region and question? Does it yield a different answer? How should we interpret counterfactual AMIP simulations in that context? What model resolution do we need to have more confidence in models?
- Is there any way to make sense of the wide range of GCM outputs? P has been mostly flat, yet the models would lead us to believe that P should be increasing. Most models can not reproduce the current drought at any point in the 21st century. What should we make of this?

Downscaling issues

- Is precipitation downscaling accurate enough for this purpose?
- My biggest concern re precipitation effects is that the USBR has traditionally used quantile mapping of precipitation fields in their downscaling, and this clearly and artificially biases the future projected precipitation changes towards wetter conditions. QM has the same wettening effect on streamflow. If the river's managers are using wet-biased projections, that is a big problem.

Other Drivers of Long-Term Change

Vegetation change (both in response to climate and to disturbance)

- Should we already have seen an effect of CO_2 fertilization on water use efficiency in the basin?
- Vegetation change. What is currently happening, what might happen?
- Beetle kill, dust, windiness
- By how much does runoff fraction for the UCRB change as vegetation changes, e.g. longer growing seasons vs increased water-use efficiency), changes in forest species and plant types as warming continues?
- I get concerned when projections show a very different and warmer climate, but our land surface models act as though the same vegetation is going to be there, transpiring water at the same rate as they would on a hot day today. This simulation of static land cover seems highly unjustified. How much will wholesale changes to the future land cover alter our runoff projections? Consider that roughly 85% of precipitation in the UCRB is lost to ET; even a small change in future ET will have a huge effect on runoff, which is just a small residual term by comparison.

Dust/aerosol

- Maybe it's just because I haven't been following the literature closely, but dust or other dark aerosol deposition on snow is a concern I have. Sure, it would make the snow melt faster and that runoff would be captured by the reservoirs, but does it also increase sublimation so that the water is entirely lost from the Colorado River system?
- Role of dust on snow. Strongly emphasized by the people who study it, less so by others. Still, seems important, but we lack lots of things: physical understanding, spatial observations, model parameterizations, etc. Same goes for vegetation changes due to land use change, fires, and beetle.

Other Questions

Processes

- ENSO effects (more 'bang for the buck' with El Niño, all else being equal?)
- Natural variability in predictability
- Do we need to understand the physical processes that drive these sensitivities? Could sensitivities change with climate so fast that historical values are irrelevant?

Models/LSM

- Do land surface models - both within GCMs and stand alone ones (e.g. VIC) sufficiently represent ET processes such that modeling can be used to answer these questions - models are not perfect e.g. simulation with ET > PET ... raising concerns about their fidelity.
- Can we better characterize the uncertainty in UCRB runoff change resulting from different treatments in land surfaces, than the current approach of using off-line LSM experiments?
- Is downscaling (e.g. using off-line LSMs with BCSD GCM-scaled T/P changes) a robust approach to determine UCRB climate response to future change?

Models/GCM

- I'm hoping we can derive recommendations that go a little beyond the Colorado River and talk more generally about challenges in hydroclimate projections, from how to use (or not use) GCMs to future observational needs and model development/evaluation needs.
- What can be done, using the combination of observations and models, to better tease out real-world temperature and precipitation responses?

- Does CMIP, in aggregate, misrepresent the climate change in Pcpn in the UCRB, owing to biases in simulated ocean warming patterns?

Management

- How does this work align with water management needs? What additional information would be most useful?

Appendix C: Breakout group transcripts

Group 1 Consensus

1. Uncertainties in obs P over the past 100 yrs make it difficult to say what the century P-forced runoff trend has actually been
2. Temperature in the UCRB has warmed at least 1 degree C in the last century
3. P is the biggest driver of Colorado River flow changes over the last 100 yrs.
4. Different obs station selection and processing can affect the meteorological trends inferred from our obs data sets.
5. Seasonality of P is a ke aspect of determining net runoff in the UCRB
6. Spatial differences inside the UCRB are extremely important – averages over the basin need to be supplemented with a finer spatial scale examination
7. Warming over the recent decades has likely decreased Colorado R flos the there are relatively wide uncertainties on how much.
8. Temperature and precipitation feedbacks (dry -> warm) and uncertain model parameterizations limit our ability to disentangle how much obs T vs P has forced historical changes.

Group 2 Consensus

1. Temperatures have warmed
2. Ensemble averages get the waring "right" (within 1/10 degree C) [of what? CMIP?]
3. Cool season precip. Is main driver of flow .
4. "Jury is out on precip trends
5. Flow trends depend on the endpoints
6. Temperature influences runoff efficiency
7. Sensitivity depends on timescale

Group 3 consensus

1. Annual surface temperature has risen on the river over roughly the last century (~1 C to 2C)
2. Neither the magnitude nor sign of P change on the river over the last century are well known. Magnitude of true change < +/-20%
3. Annual Streamflow at Lees Ferry has declined over the last century (-10% to -15 %)

4. The "Millennium Drought" post-2000 is the most sustained drought in terms of consecutive (cumulative) low flows during the instrumental period [Note I inserted "cumulative" based on discussion at the time]
5. Paleo=-data (1400-present) reveal several sustained droughts analogous to the Millennium Drought [see C. Woodhouse slides for precise definitions]
6. Temperature has contributed to low flow during the Millennium drought, but with unknown magnitude

Group 4 consensus

1. There has been a warming trend that is not linear of about 1 degree C. (i.e. warming from 1900 – 1940, flat from 1940 -late 1970s, warming to present) Does the nonlinearity matter?
2. Winter precipitation is more important for Water Year flow. Attempted rationale"
    a. Theory – less evap than for summer precipitation bc of colder temps, less radiation
    b. Soil moisture is higher, or frozen so more runs off
    c. Residence time in soils shorter during rapid snowmelt) runoff
    d. Correlation of P_winter with Q_qy (i.e. observational analysis
    e. Hydro model sensitivity studies
3. Data are less reliable before 1950, 1979 (beginning of Snotel). Temperature anomalies better represented than precipitation because they are larger scale. But CO-OP stations may have declined in mountainous areas, so can't be assumed to be monotonically increasing
4. CMIP5 models really are wetter than CMIP3, but large biases (typically wet biases) remain
5. Temperature sensitivity range broadens on both ends from Vano et al, 2014.

Group 1 Research Gaps

1. During the period of high flow (!1900-1920) was there comparably high precipitation?
    a. Mining precipitation records
    b. Tree ring data
    c. Dynamical/statistical downscaling of 20CR/ECR reanalyses
    d. Bayesian Inferred precipitation from streamflow
2. Better characterization of the physics controlling the Millennium Drought: Drought or aridification?
    a. Role of groundwater

   b. Role of Temperature and precipitation
   c. Reconciling different sensitivities
   d. Role of seasonality (control 1980's, compare cold/dry vs hot/dry (2000-2010)
      i. LSM MIP
      ii. Couplet atmos hydro MIP (i.e. WRF Hydro)
      iii. GCM AMIP
3. Assessment of predictability of the annual flow of the river on different timescales
   a. AMIP
   b. Initialization of forecast models (decadal prediction studies)
   c. Statistical/dynamical systems analyses
4. Consistency of data analysis methods and metrics

Group 2 Research Gaps

1. Can we get paleo estimates of elasticity, efficiency, and sensitivity?
2. How do elasticity, efficiency, and sensitivity vary by sub-basin?
3. How sensitive re state-of-the-art methods for estimating elasticity, efficiency and sensitivity to observational uncertainties.
4. Model consistency, for example, in terms of variance (for example paleo only explains part of the variance)?

Group 3 Research Gaps

1. Which seasons are really key to historical and future changes in runoff?
2. What are the best GCM metrics for evaluating UCRB credibility?
3. Why is there a big difference between historically based estimates of T, P sensitivity and model based estimates?
4. How much are we being mislead by considering all of the non-precip changes to be due to temperatures, ignoring dusd, dead trees, other unknown forcings?
5. Need better uncertainty estimates on gridded dataset -- where are the uncertainties coming from and can the analysis of currently unincluded data sources shed light?
6. What spatial resolution do we really need to model the UCRB?

Group 4 research gaps

1. Using historical record to constrain/inform model application.
   a. Understand cause for the wider range

    b. Define “emergent” metrics (i.e. models not tuned to these metrics) for evaluation and constraining ESMs GCMs, LSMs etc. Sensitivity and elasticity are examples of emergent metrics.
    c. Run all LSMs in the GCMs (e.g. in the CMIP6 GCMs) offline with historical observations. [or make such data available – will one of the CMIP6 MIPs do this?]
2. Disentangling multiple forcing factors for historical flow change
    a. Dust, vegetation change [dead trees], land use, temperature and precipitation
    b. Seasonality
    c. Consider a suite of experiments with the NCAR modeling suite?
3. More rigorous uncertainty estimate of gridded data