Appendix F

Approach to Hydrologic Uncertainty

This page intentionally left blank.

Contents

Table

Figures

This page intentionally left blank.

Appendix F. Approach to Hydrologic Uncertainty

F.1 Introduction & Purpose

Future hydrology, the primary driver of how each alternative performs, is deeply uncertain over the medium-to-long-term (Smith et al., 2022). Therefore, alternatives in the Draft Environmental Impact Statement (DEIS) are evaluated according to their ability to maintain resource-specific performance goals across a broad range of streamflow conditions (i.e., robustness analysis) and the streamflow conditions in which each alternative is likely to fail those goals (i.e., vulnerability analysis) (McPhail et al., 2018; Bonham et al., 2024; Bonham, Kasprzyk and Zagona, 2025). This approach, a type of Decision Making Under Deep Uncertainty (DMDU; see **Appendix D**, 'DMDU Overview and Approach') (Marchau et al., 2019), requires a systematically selected set of streamflow traces to use as inputs to the resource modeling.

The modeling in the DEIS uses 400 streamflow traces selected from a combination of five streamflow ensembles and a statistical subsampling approach. All traces from the Stress Test, Coupled Model Intercomparison Project (CMIP) Phase 5 – Localized Constructed Analogs (CMIP5-LOCA), and Drying with Variability ensembles are included (200 traces), and 200 more traces were subsampled from the Paleo Drought Resampled and CMIP3 Paleo-Conditioned ensembles using the Kennard Stone algorithm (totaling 400 traces).

The purpose of this appendix is to describe the five streamflow ensembles (**Section F.2**), explain the subsampling approach (**Section F.3**), and discuss the resulting set of 400 traces (**Section F.4**). **Section F.5** provides a summary of previous sections.

F.2 Streamflow Ensembles

F.2.1 Stress Test

The Stress Test ensemble provides continuity with many recent studies (e.g., 2019 Drought Contingency Plan) and routine long-term projections. The ensemble was created by resampling the 1988-2023[1] subset of the historical natural flow record using the index sequential method (Ouarda, Labadie and Fontane, 1997) and focusing on recent hydrology. This recent period has a 11% drier average flow than the full historical record (1906-2023). Use of the Stress Test ensemble is supported by multiple research studies that identified a shifting temperature trend in the Colorado River Basin in the late 1980s that affected runoff efficiency and resulted in lower average flows for

[1] 2021-2023 natural flow data are estimates based on the April 2024 24-Month Study (https://www.usbr.gov/lc/region/g4000/NaturalFlow/LFnatFlow1906-2024.2024.4.22.xlsx)

the same amount of precipitation (Woodhouse et al., 2016; McCabe et al., 2017; Udall and Overpeck, 2017).

F.2.2 CMIP5-LOCA

The CMIP5-LOCA ensemble uses the best available projections from global climate models that are available at the scale necessary for input to CRSS. The ensemble uses Upper Basin flows from a project completed by Reclamation's Technical Services Center for the Upper and Lower Colorado Regions that relies on data from Vano et al. (2020) (Reclamation, 2025). These flows were created by running LOCA downscaled climate projections of temperature and precipitation through the Variable Infiltration Capacity (VIC) hydrological model. No secondary streamflow bias correction was applied to the VIC flows. The ensemble includes 64 traces: 32 traces derived from climate projections for representative concentration pathway (RCP) 4.5 and 32 traces from RCP 8.5. Natural flows for Lower Basin sites were determined by applying the k-Nearest Neighbors (kNN; Prairie et al., 2008; Nowak et al., 2010) approach, which selects flows from the historical year whose Lees Ferry flows most closely match those of the projected year.

F.2.3 "Drying with Variability"

The "Drying with Variability" ensemble was developed in collaboration with Utah State University. Through a multiyear project that began in 2021, researchers performed in-depth analysis on the ensembles that were available for Reclamation's long-term modeling to identify any gaps in statistical characteristics that could (a) be important for testing the system, and (b) capture aspects of the best available information from different sources (Salehabadi et al., 2022). Analysis showed that none of the available ensembles included a long-term drying trend while maintaining or increasing variability in flow magnitudes.

As part of the project, an ensemble was developed that met these criteria using a statistical approach to combine information from multiple established data sources (Salehabadi et al., 2025). Specifically, the ensemble combines the following information:

- Observed Lees Ferry annual flows from 1931 to 2020 with wet and dry sequencing from the paleo-reconstructed record;
- An extrapolation of the observed temperature trend with an estimate of how runoff efficiency[2] is affected by warmer temperatures;
- Projections from multiple generations of global climate models that indicate that precipitation variability could increase in the future.

Through nonparametric resampling of observed flows and paleo-reconstructed sequencing, application of the long-term declining trend, and a new method for transforming the flow variability while preserving the trend, an ensemble was generated that filled the statistical and narrative gap.

[2] Runoff efficiency is a measure of the amount of streamflow produced for a given amount of precipitation. Runoff efficiency has been shown to decline as temperatures increase due to increased evaporation, drying of soils, and increased evapotranspiration from plants.

To prepare this ensemble for use in CRSS, annual Lees Ferry flows are disaggregated to all 29 natural flow locations using the kNN method (Prairie et al., 2008; Nowak et al., 2010).

The Drying-with-Variability is used as the 'reference hydrology' included with every vulnerability analysis in **Volume III**. The reason this ensemble was chosen as the reference hydrology is because it helps the reader interpret the likelihood of undesirable outcomes (e.g., Lake Mead dead pool-related reductions) if the current downward trend continues and/or the frequency and duration of droughts become more severe. Given the significant decline in reservoir storage that has occurred from 2000-2025, long-term planning in the Basin must wrestle with the potential for such a future.

F.2.4 Paleo Drought Resampled (subsampled)

The Paleo Drought Resampled ensemble is a severe drought scenario developed by Salehabadi et al. (2022). The ensemble randomly resamples Lees Ferry flows from the 16th century drought (1576-1600) from Meko et al. (2017) paleo reconstruction to create 100 traces that include unique sequences of flows from this period. The ensemble was subsampled to select 50 traces that fill in patterns missing from the three complete ensembles, as described in the next section. Annual Lees Ferry flows are disaggregated to all 29 natural flow locations using the kNN method (Prairie et al., 2008; Nowak et al., 2010).

F.2.5 CMIP3 Paleo-Conditioned (subsampled)

The CMIP3 Paleo-Conditioned ensemble is developed using two valuable data sources and aligns with a recommendation from Colorado River Basin Climate and Hydrology State of the Science Report (Lukas and Payton, 2020). The ensemble was created using the non-parametric conditioning (NPC) method (Prairie et al., 2008) with the Meko et al. (2017) paleo record to define wet-dry sequencing and sampled magnitudes from annual CMIP3 flows. This NPC sampling method was performed on each of the 112 CMIP3 streamflow projections (2024-2060) to create 50 traces for each projection, resulting in 5,600 traces for all 29 natural flow sites. The ensemble was subsampled to select 150 traces that fill gaps in patterns and characteristics, as described in the next section.

F.3 Subsampling Approach

The subsampling was performed using a 3-step procedure and the Kennard Stone algorithm (Kennard and Stone, 1969), which, in this application, selects streamflow traces from the Paleo Drought Resampled and CMIP3 Paleo-Conditioned ensembles with statistics not well-represented by traces in the Stress Test, CMIP5-LOCA, or Drying with Variability ensembles.

This section is organized as follows. **Section F.3.1** describes the Kennard Stone algorithm. **Section F.3.2** explains the streamflow statistics used as inputs to Kennard Stone. **Section F.3.3** describes the 3-step sampling procedure.

F.3.1 Subsampling algorithm: Kennard Stone

Given an existing sample of n data points, the Kennard Stone algorithm selects k additional points (for a total of $n + k$ points) from population P such that the k points are maximally different from the original *n* and from each other[3]. The algorithm uses an iterative approach; it first identifies k_1 (the first data point to be selected) as the point that maximizes the distance from all *n* points, where distance is calculated using statistics associated with each point and a selected distance metric. Then, k_2 is selected to maximize the distance from all *n* points and k_1.This process continues until all *k* points have been selected (Kennard and Stone, 1969; Stevens and Ramirez-Lopez, 2020).

In this application, each streamflow trace is a data point. The existing sample of *n* data points are the traces in the Stress Test, CMIP5-LOCA, and Drying with Variability ensembles (*n* = 36 traces [Stress Test] + 64 traces [CMIP5-LOCA] + 100 traces [Drying with Variability] = 200 traces). *k* = 200 so the total number of traces is 400. Mahalanobis distance (Mahalanobis, 1936; Brereton, 2015) was chosen for the distance metric to account for correlation between streamflow statistics, which are described in the next section.

F.3.2 Streamflow statistics

The streamflow statistics (**Table F-1**) are calculated using 2027-2056 calendar year volumes of natural flow at Lees Ferry, Arizona. The statistics describe dry, wet, and normal flow conditions over various durations (1-, 2-, 5-, 10-, and 20-year windows) to ensure the set of 400 traces cover a broad range of hydrologic conditions. The statistics are calculated for each trace of every ensemble, which are then used in the 3-step application of Kennard Stone described in **Section F.3.3**.

Table F-1
Streamflow statistics used with the Kennard Stone sampling algorithm

Streamflow Statistic	Category
Minimum 1-year natural flow volume	Dry conditions
Average Lees Ferry natural flow during the driest 5-year period	
Average Lees Ferry natural flow during the driest 10-year period	
Average Lees Ferry natural flow during the driest 20-year period	
Maximum 1-year natural flow volume	Wet conditions
Average Lees Ferry natural flow during the wettest 5-year period	
Average Lees Ferry natural flow during the wettest 10-year period	
Average Lees Ferry natural flow during the wettest 20-year period	
Median of the rolling 2-year rolling average natural flow	Normal conditions
Median of the rolling 5-year rolling average natural flow	
Median of the rolling 10-year rolling average natural flow	
Median of the rolling 20-year rolling average natural flow	
2027-2056 Average Lees Ferry natural flow	Long-term average

F.3.3 3-Step Sampling Approach

The following 3 steps were used to subsample a total of 200 traces from the Paleo Drought Resampled and CMIP3 Paleo-Conditioned ensembles.

[3] If *n* = 0 (i.e., zero data points have been selected to start with), then k_1 and k_2 are selected such that they maximize the pairwise distance between any two points in the population.

1. Sample 50 traces from the Paleo Drought Resampled ensemble using Kennard Stone. The 50 traces capture the variability present in this ensemble while saving computing time compared to keeping all 100 traces. This brings the total number of traces to 250.

2. Remove very wet traces from CMIP3 Paleo-Conditioned ensemble using filter criteria. This dataset contains a limited number of very wet traces because annual flows originate from the CMIP3 hydrology dataset. The goal was to systematically identify and remove these traces to avoid 'shifting' the distribution of the final 400 traces wetter due to dubious high flows. To do so, the CMIP3 Paleo-Conditioned traces were compared to the maximum values for select statistics from the combined Stress Test, Paleo Drought Resampled, and Drying with Variability ensembles. If a given trace was wetter according to any of the selected statistics, then it was removed before step 3. The criteria can be expressed mathematically as:

   - maximum annual flow < 26.45 MAF & maximum 2-Year average flow < 26.11 MAF/year & maximum 5-year average flow < 24.61 MAF/year & 2027-2056 average flow < 16.47 MAF/year

   996 traces were removed because they failed one or more of the criteria, leaving 4,604 in traces in the CMIP3 Paleo-Conditioned ensemble before moving to step 3.

3. Sample 150 traces from the pre-filtered CMIP3 Paleo-Conditioned ensemble. This brings the total number of traces to 400.

F.4 Results

Figure F-1 shows the distribution of trace average natural flow, grouped by ensemble (x-axis) and by time period (panels). All 400 traces are included in the boxplot labeled 'Combined,' while the remaining boxplots show traces grouped by ensemble. The time period averages are shown because the correspond to the periods used in the robustness analyses found in **Volume I – Chapter 3** and **Volume III**. For comparison, the dashed lines labelled 14.6, 13.0, and 12.4 show the historical average natural flow in MAF from 1906-2024 (full historical record), 1988-2024 (Stress Test period), and 2000-2024 (Millenium Drought), respectively.

The combined set of 400 traces covers the full range of average natural flow present in any individual ensemble while providing a nearly continuous sampling from wet to dry extremes. This is important because this DEIS evaluates critical performance goals that can fail under dry conditions (e.g. dead pool related delivery reductions) and wet conditions (e.g. use of the Glen Canyon Dam spillway). Across the Full Modeling Period (2027-2060), average natural flow ranges from about 8 to 20 MAF. Comparing the Combined boxplot to historical averages, about 50% of the traces are drier than the 1988-2024 average, over 25% of the traces are drier than the 2000-2024 average, and about 10% of the traces have an average below 11 MAF/year. In the subperiods (2027-2039, 2040-2049, and 2050-2060), the 25th, 50th and 75th percentiles are similar to the Full Modeling Period while the tails of the distribution are larger in both the wet and dry direction. This is due to the averages being calculated over 10 to 13 years in the subperiods compared to the 34 years in the Full Modeling Period, which dampens the effect of extremely wet and dry years on the reported averages.

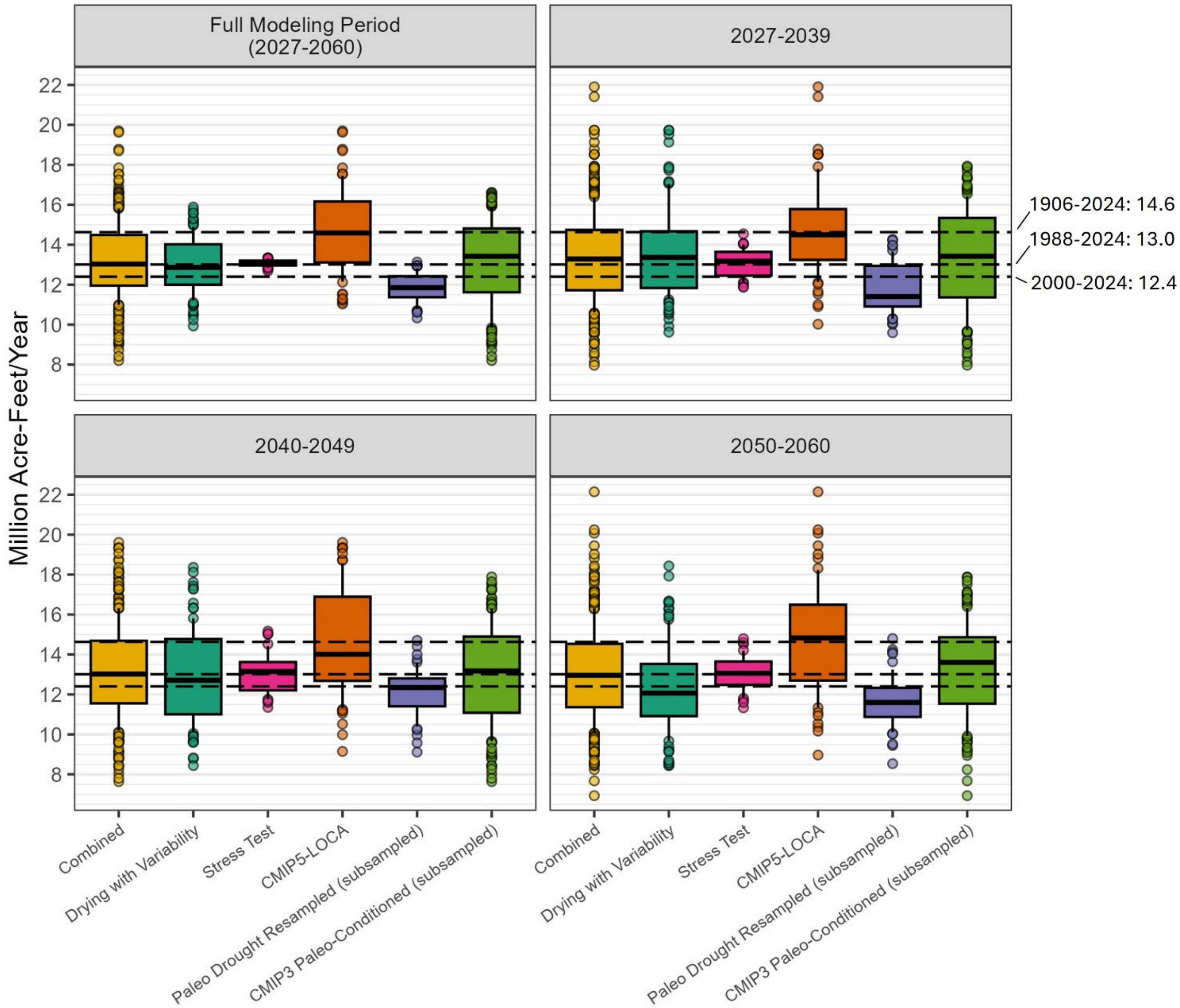
Figure F-2 shows the distribution of the driest 1-, 2-, 5-, 10-, and 20-year average volumes in each trace, grouped by ensemble. The dashed line shows the minimum volumes for each period observed during the historical period (1906-2024). In the Combined ensemble, about 25% of traces have a driest 1-year volume less than the historical minimum. For the 2- to 20-year trace averages, 50% or more of traces in the Combined ensemble are drier than the historical minimum.

Figure F-3 shows the distribution of the wettest 1-, 2-, 5-, 10-, and 20-year average volumes in each trace, grouped by ensemble. The dashed line shows the maximum volumes for each period observed during the historical period (1906-2024). Across 1-to 20-year trace averages, nearly 75% to over 90% of traces in the Combined ensemble have values less than the historical maximums.

Considering the wettest 1-year volumes, most traces above the 90th percentile in the Combined ensemble are from the CMIP5-LOCA ensemble. This is also true for the 2- to 20-year volumes, but with some of the very wet traces originating from the Drying with Variability and CMIP3 Paleo-Conditioned ensembles.

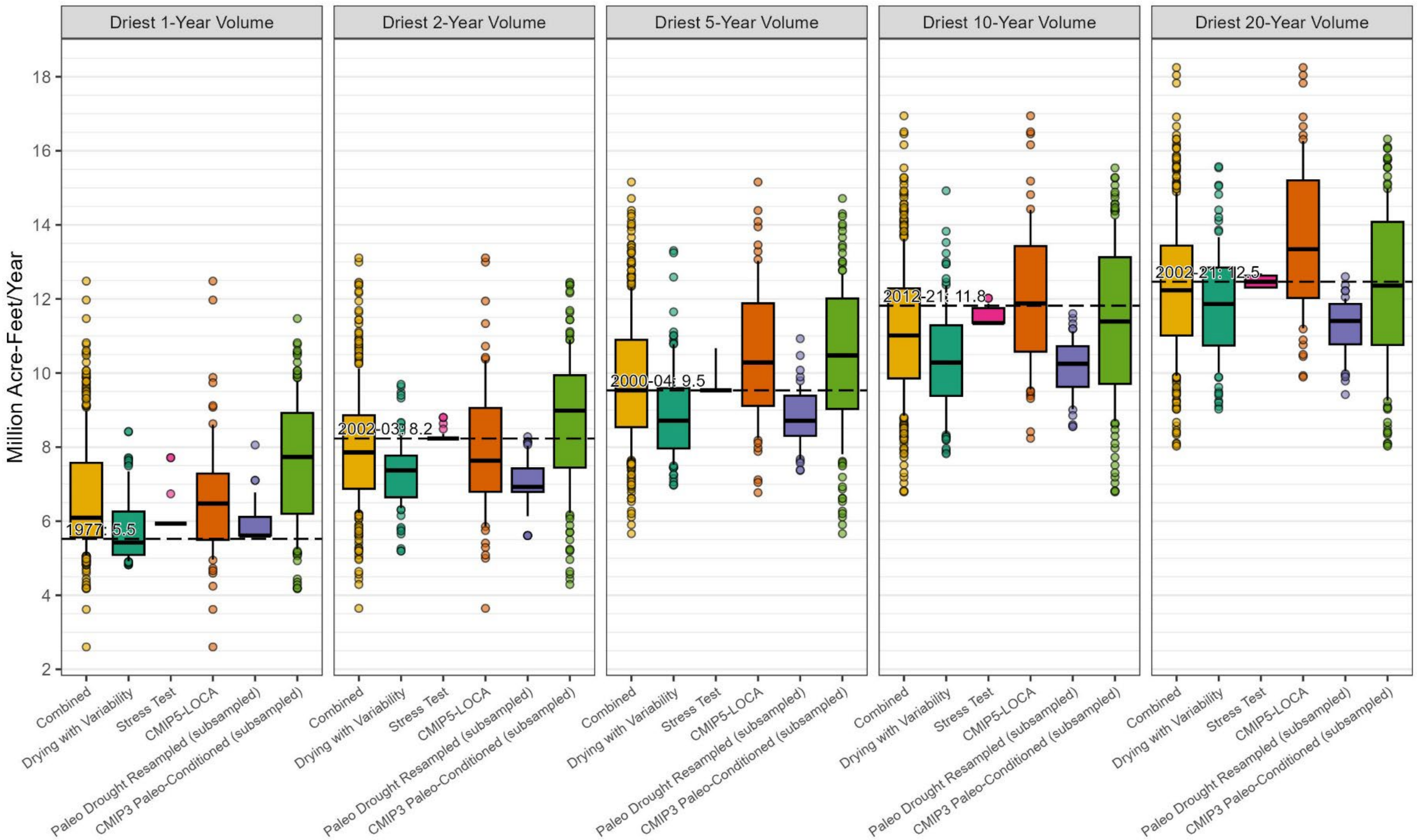
The distribution of average natural flow for the Drying with Variability ensemble, which is used as the reference hydrology in the vulnerability analysis figures in **Volume III**, demonstrates a storyline of declining streamflow over time. Over the full modeling horizon, the median average natural flow is 12.9 MAF/year (slightly less than the median of the combined ensemble, 13.0 MAF/year). In each subsequent subperiod, however, the median average natural flow decreases from 13.4 to 12.7 to 12.1 MAF/year.

Figure F-1
Distribution of average water year natural flow at Lees Ferry, grouped by ensemble (x-axis) and by time period (panels).



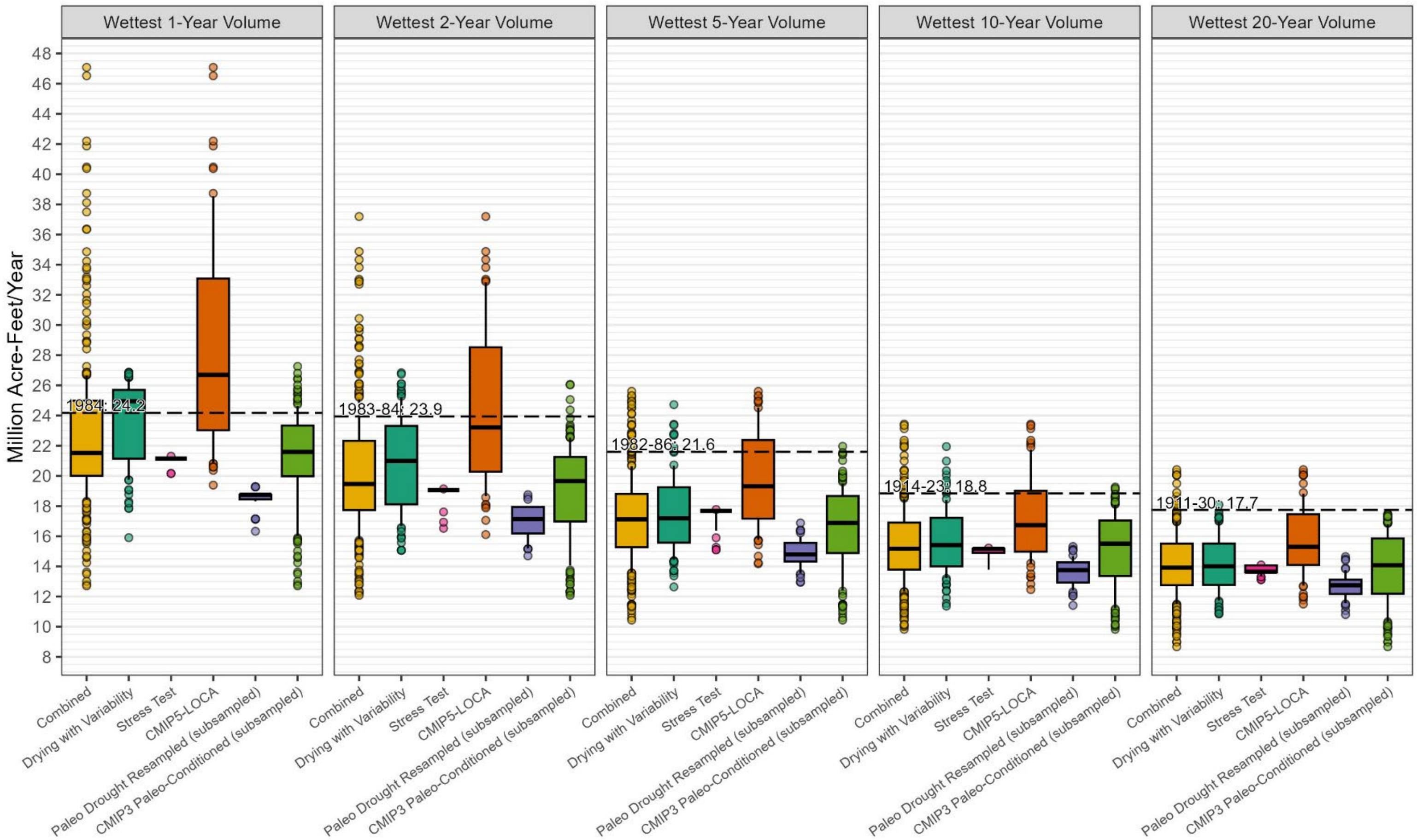
Note: The dashed lines labelled 14.6, 13.0, and 12.4 indicate the historical average natural flow in million acre-feet from 1906-2024 (full historical record), 1988-2024 (Stress Test period), and 2000-2024 (Millenium Drought), respectively.

Figure F-2
Distribution of the driest 1-, 2-, 5-, 10- and 20-Water Year Volumes (columns) in each trace, grouped by ensemble.



Note: Dashed lines show the minimum volumes for each period observed during the historical period (1906-2024).

Figure F-3
Distribution of the wettest 1-, 2-, 5-, 10- and 20-Water Year Volumes (columns) in each trace, grouped by ensemble.



Note: Dashed lines show the maximum volumes for each period observed during the historical period (1906-2024).

F.5 Summary

Due to the deeply uncertain nature of future hydrology in the Colorado River Basin, DMDU is the analytical framework used in this DEIS. To reliably evaluate the robustness and vulnerability of the alternatives, it is essential to model them with a systematically selected set of streamflow traces. This DEIS uses 400 total traces from the following ensembles:

- Stress Test (36 traces)
- CMIP5-LOCA (64 traces)
- Drying with Variability (100 traces)
- Paleo Drought Resampled (subsampled, 50 traces)
- CMIP3 Paleo-Conditioned (subsampled, 150 traces)

All traces from Stress Test, CMIP5-LOCA, and Drying with Variability ensembles were used because they provide continuity with previous analyses and useful storylines about how future streamflow could decline. Combined, these ensembles span a wide range of plausible streamflow conditions but leave gaps in the types of conditions being evaluated. To fill in the gaps, a subsample of traces was selected from the Paleo Drought Resampled and CMIP3 Paleo-Conditioned ensembles using the Kennard Stone algorithm. The result is a combined set of 400 traces that cover a broad range of plausible conditions with minimal gaps in the types of conditions being tested.

F.6 References

Bonham, N., J. Kasprzyk, E. Zagona, and R. Smith. 2024. "Interactive and Multimetric Robustness Tradeoffs in the Colorado River Basin," *Journal of Water Resources Planning and Management* 150(3):05023025. Internet website: https://doi.org/10.1061/JWRMD5.WRENG-6199.

Bonham, N., J. Kasprzyk, and E. Zagona. 2025. "Taxonomy of purposes, methods, and recommendations for vulnerability analysis," *Environmental Modelling & Software* 183:106269. Internet website: https://doi.org/10.1016/j.envsoft.2024.106269.

Brereton, R. G. 2015. "The Mahalanobis distance and its relationship to principal component scores: The Mahalanobis distance and PCA," *Journal of Chemometrics* 29(3):143–145. Internet website: https://doi.org/10.1002/cem.2692.

Bureau of Reclamation (Reclamation). 2025. Colorado River Hydrologic Analysis for Post-2026 Operations Planning. Technical Memorandum Number. ENV-2024-050.

Kennard, R. W. and L. A. Stone. 1969. "Computer Aided Design of Experiments," *Technometrics* 11(1):137–148. Internet website: https://doi.org/10.1080/00401706.1969.10490666.

Lukas, J. and E. Payton. 2020. Colorado River Basin Climate and Hydrology: State of the Science. Boulder, CO: Western Water Assessment. Internet website: https://scholar.colorado.edu/concern/reports/8w32r663z.

Mahalanobis, P. 1936. "On the generalized distance in statistics." *Proceedings of the National Institute of Science in India* 2:49–55.

Marchau, V. A. W. J., W. E. Walker, P. J. T. M. Bloemen, and S. W. Popper. 2019. *Decision Making under Deep Uncertainty: From Theory to Practice*. Cham: Springer International Publishing. Internet website: https://doi.org/10.1007/978-3-030-05252-2.

McCabe, G. J. D. M. Wolock, G. T. Pederson, C. A. Woodhouse, and S. McAfee. 2017. "Evidence that Recent Warming is Reducing Upper Colorado River Flows," *Earth Interactions* 21(10):1–14. Internet website: https://doi.org/10.1175/EI-D-17-0007.1.

McPhail, C., H. R. Maier, J. H. Kwakkel, M. Giuliani, A. Castelletti, and S. Westra. 2018. "Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results?," *Earth's Future* 6(2):169–191. Internet website: https://doi.org/10.1002/2017EF000649.

Meko, D., C. Woodhouse, and E. Bigio. 2017. University of Arizona Southern California Tree-Ring Study. Agreement 4600011071. Internet website: https://data.ca.gov/dataset/paleo-dendrochronological-tree-ring-hydroclimatic-reconstructions-for-northern-and-southern-cal/resource/3b654c43-d5b1-43db-b5ef-71bb46721990.

Nowak, K., J. Prairie, B. Rajagopalan, and U. Lall. 2010. "A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow," *Water Resources Research* 46(8). Internet website: https://doi.org/10.1029/2009WR008530.

Ouarda, T. B. M. J., J. W. Labadie, and D. G. Fontane. 1997. "Indexed sequential hydrologic modeling for hydropower capacity estimation," *Journal of the American Water Resources Association* 33(6):1337–1349. Internet website: https://doi.org/10.1111/j.1752-1688.1997.tb03557.x.

Prairie, J., K. Nowak, B. Rajagopalan, U. Lall, and T. Fulp. 2008. "A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data," *Water Resources Research* 44(6). Internet website: https://doi.org/10.1029/2007WR006684.

Salehabadi, H., D. G. Tarboton, B. Udall, K. G. Wheeler, and J. C. Schmidt 2022. "An Assessment of Potential Severe Droughts in the Colorado River Basin," *Journal of the American Water Resources Association* 58(6):1053-1075. Internet website: https://doi.org/10.1111/1752-1688.13061.

Salehabadi, H., D. G. Tarboton, K. Wheeler, J. Prairie, R. Smith, and S. Baker. 2025. "Developing Storylines of Plausible Future Streamflow and Generating a New Warming-Driven Declining Streamflow Ensemble: Colorado River Case Study," *Water Resources Research* 61(1):e2024WR038618. Internet website: https://doi.org/10.1029/2024WR038618.

Smith, R., E. Zagona, J. Kasprzyk, N. Bonham, E. Alexander, A. Butler, J. Prairie, and C. Jerla. 2022. "Decision Science Can Help Address the Challenges of Long-Term Planning in the Colorado River Basin," *Journal of the American Water Resources Association* 58(5):735–745. Internet website: https://doi.org/10.1111/1752-1688.12985.

Stevens, A. and L. Ramirez-Lopez. 2020. An introduction to the prospectr package. R package Vignette. Internet website: https://cran.r-project.org/web/packages/prospectr/prospectr.pdf.

Udall, B. and J. Overpeck. 2017. "The twenty-first century Colorado River hot drought and implications for the future: Colorado River flow loss," *Water Resources Research* 53(3):2404–2418. Internet website: https://doi.org/10.1002/2016WR019638.

Woodhouse, C. A., G. T. Pederson, K. Morino, S. A. McAfee, and G. J. McCabe. 2016. "Increasing influence of air temperature on upper Colorado River streamflow: temperature and Colorado streamflow," *Geophysical Research Letters* 43(5):2174–2181. Internet website: https://doi.org/10.1002/2015GL067613.

Vano, J., J. Hamman, E. Gutmann, A. Wood, N. Mizukami, M. Clark, D. W. Pierce, D. R. Cayan, C. Wobus, K. Nowak, and J. Arnold. 2020. Comparing Downscaled LOCA and BCSD CMIP5 Climate and Hydrology Projections - Release of Downscaled LOCA CMIP5 Hydrology.