Peer-review
"Assessing decadal to centennial scale nonstationary variability in meteorological drought trends"
General comments

Reviewer #1
This manuscript presents a clear description of the research gap, models, and analysis. The improvements in the document after the first peer-review round are substantial, providing a clear perspective on the research. Small corrections should be addressed to improve the manuscript.

> The authors really appreciate your valuable comments to substantially improve the article. We did our best to respond to each of your helpful comments. Our responses are below.

Minor comments
Line 172. Please explain why the 94th percentile threshold and 6th percentile represent the precipitation associated with dry anomaly, and wet anomaly for each location?

> These percentiles correspond to SPI values of -1.5, 0, and 1.5 (see Line 150). These thresholds are commonly used in drought and pluvial studies to represent the precipitation associated with severe dry anomaly, typical, and wet conditions (Heim, 2002). For example, McKee et al. (1993) used -1.5 as a threshold between "Moderate" and "Severe" droughts. This is the original source for the SPI and the threshold has propagated into many later publications. The US Drought Monitor uses a similar threshold of -1.6.
>
> We have added a sentence to explain in the main text Line 150:
> These thresholds are commonly used in drought and pluvial studies to represent precipitation associated with dry anomaly, typical, and wet conditions (Heim, 2002), and a similar thresholds (5th percentile) is used by the US Drought Monitor (Svoboda et al. 2002).

Line 179. Please explain the insert K historical 13-Months precipitation sequences.

> We have added a sentence to explain 13months sequence.
> Lines 159 – 160:
> The 13-month sequence is considered a single downscaling unit containing three known NASPA values across a year (Figure 1).

Line 312. This correspond to the K historical 13-months precipitation sequence? If this is the case should be explained better before.

> Thank you for the suggestion. We made a clarification in lines 158-159. We believe this is smoothly connected to the detail explanation in lines 166-168.

Reviewer #2

The authors have addressed all my comments and I find the paper can be accepted for publication. Remaining comments (below) can be addressed in a final minor revision.

> The authors greatly appreciate for your time and thorough comments. We believe our manuscript has benefited from your helpful comments. Our responses are below.

L212: I think Gamma distribution parameters are not mean and shape, but shape and scale (pls revise throughout the text)

> The GAM model used by the 'mgcv' package estimates mean and shape parameters for the gamma distribution, instead of the shape and scale. The scale parameter can be then estimated as scale = mean/shape. We have modified equations 1 and 2 to clarify this.
>
> Estimating mean rather than scale is because the underlying model began as a more traditional gamma regression Generalized Linear Model (GLM), where the mean is estimated and the shape (or dispersion) parameter is estimated, but fixed. The mgcv package expands this to allow regression for the shape parameter. In addition to being more easily interpretable, a focus on mean and shape makes the resulting parameter estimates more robust.
>
> We have also added following statement in Lines 211 - 213:
>
> The Gamma distribution is typically prescribed by shape and scale parameters (α and θ, respectively), but our approach follows Wood (2006), instead estimating the mean and shape parameters (Eqs. 2-3). The scale parameter can then be estimated from the mean and shape (Eq. 1).

L281: I do not see a generally good downscaling skill in season DJF in Fig. 4 (where there is also no tree growth in most climates), but in JFM-JJA. In addition, some stations show a good downscaling skill also in winter – can it be explained by the climate?

> Reconstruction skill in the original NASPA is often better during winter/spring (DJFMA) precipitation than summer (MJJ) for the southwestern US (Table 1 and Stahle et al. 2020). For example, Los Angeles, CA, Rodeo, NM, and Phoenix, AZ. This is because winters are generally warmer, the DJFMA period includes early spring, and winter rainfall generally sets up soil moisture for tree growth later in the year. Summers tend to be quite arid, making discriminating between precipitation anomalies more difficult for tree-rings, which has a warmer winter climate and often quite generally shows good skills in winter (Stahle et al., 2020).
>
> We have modified Lines 291- 298 to better clarify this:
>
> There are a few exceptions showing the best skill during winter (DJF) and poor skill, large nMAE, during early summer (MJJ, Figure 4). This occurs in only in the southwestern US (e.g. Los, CA, Phx, AZ and Roe, NM) where the underlying NASPA shows better initial reconstruction skill during their relatively mild winters (Table 1) and less skill in summer. For the dsNASPA, the seemingly large errors in nMAE during MJJ are primarily due to extremely small values in the denominator of the nMAE due to very low precipitation, combined with not capturing infrequent large rainfall events.

Fig. 5 Capture: Think there are no 'mean parameter' estimates shown. Rather estimates of the long-term trend in each data set.

> We have modified the figure caption as follows:
>
> Original: Mean parameter estimates for each dataset are represented…
>
> Fixed: The modeled long-term trends from each dataset are represented…

L321: Bias in the mean parameter? (Gamma has no mean parameter, see above). Check also for shape parameter…

As mentioned above, the GAM model in 'mgcv' package directly fits the mean and shape parameters, whereas the scale parameter is a secondary estimate calculated from these two. We would like to leave the text as it is for the consistency, as the section 3.2 describes our model bias in technical context and mentions mean bias a few more times. However, as mentioned above, we have clarified this relationship between mean, shape, and scale in Eqs. 1-3, along with modifications to the text.

L344: (add at end of sentence): "... while the precipitation series are shown as raw data without bias correction for context."

Added at the end of the sentence.

Line 351:

The solid black line shows the common, long-term trend of the mean while the precipitation series are shown as raw data without bias correction for context.

L 346: 'abrupt' appears too strong for what I would observe in Fig. 6

We modified 'abrupt to 'noticeable' (line 353 now). We have also changed 'an abrupt drying trend' in line 351 to 'a drying trend'.

Fig- 6: "The yellow shaded area represents the 95% percentile for the Gamma distribution" should be reformulated. Think this is a 95% confidence interval of the long-term trend based on a Gamma distribution (or: assuming Gamma-distributed precipitation series)

Agreed. It was a precipitation range between SPI = 1.5 and -1.5. We have modified the caption as below:

The yellow shaded area represents the precipitation amount between SPI = +1.5 (upper boundary) and -1.5 (lower boundary) in the fitted Gamma distribution.

L359: Either "range between SPI = 1.5 and SPI = -1.5 ..." Or "precipitation variance"

Agreed. We changed to "The range between SPI = 1.5 and SPI = -1.5".

L. 385: One can hardly see this interpretation in the plots

Agreed. We have removed the sentence to avoid confusion.

Supplement: Check figure captions

We corrected the figure number.