Response to Reviewer #1's Comments on:
"***Subseasonal prediction of springtime Pacific-North American transport using upper-level wind forecasts***"
by J. Albers et al. (WCD paper # wcd-2020-60)

We thank Reviewer #1 for their comments, suggestions, and reference suggestions, which significantly improved the manuscript. We discuss the references that we have added below.

Reviewer wrote: *I wonder why only the spring months were examined. I believe the seasonality of the STT would limit the ability during summer and fall, but I would think the magnitude of STT during DJF is great enough to be useful. At the very least, a short discussion of the other seasons and the possibility of prediction outside of MAM should be presented.*

Our response: The reviewer is certainly correct that DJF skill (at least for TME and STT to 500 mb) is likely large enough to be useful. That said, we are specifically interested in MAM for several reasons. Since the reviewer specifically asked about STT, we will discuss this first, but we follow with an overall rationale that applies to both STT and TME.

The original motivation for our study was to examine STT to the PBL over North America, because this is the type of STT that can directly influence surface air quality. For this specific type of STT, spring is by far the most important time period, which our study is not the first to point out (see for e.g. the references we included: Fiore et al. 2003; EPA US 2006; Langford et al. 2009; Lefohn et al. 2011; note that our Fig. 1-first column also supports this).

In terms of STT to 500 hPa and TME (and water vapor transport more generally) on the other hand, the reviewer is quite right that transport skill during DJF is an interesting topic. Indeed, water vapor transport is large in magnitude (e.g., atmospheric rivers from the Pacific basin to the west coast of the US) and forecast skill perhaps at a peak during DJF. This is likely one reason why there are so many studies (too many to count) that discuss the predictability of water vapor transport during DJF. However, there are very few modern studies (if any?) that have examined STT and TME forecast skill during spring, despite the fact that this is a dynamically interesting period (see for e.g., the Breeden et al. paper we cite) and is an important season for moisture transport over the Pacific-North American region (e.g., Mundhenk et al. 2016).

Thus, because (1) STT to the PBL is so important during spring and (2) very little research has documented STT and TME forecast skill during spring, we feel that focusing on spring rather than other months is a worthwhile endeavor. That said, we agree with the reviewer that we should be more explicit when explaining our motivation for examining spring, so we have rewritten a portion of the Introduction and added some additional that expands on our rationale for focusing on MAM (see lines 31-46).

Reviewer wrote: *Lines 185-187: It is not clear how the claim of "vertically deep" shifts in the jet stream is made from the 2-D comparison. I am assuming that is because the jet climatology is derived from vertically averaged winds between 100-500 hPa. However, given the strong vertical gradients of wind around the jet stream, the wind changes don't have to be vertically deep to result in shifts to the jet. In addition, the phrase could also be interpreted to imply the jet stream location shifts vertically which is not necessarily the case.*

Our response: You are correct in assuming that the jet climatology was calculated for vertically averaged winds between 100-500 hPa. Although we did not include the results in the manuscript, we also tested our results using a "shallow jet" climatology (based on upper tropospheric winds only, which was also taken from the Koch et al. 2006 dataset). We found that our results were not dependent on the jet climatology depth. Thus, it would seem that the depth is probably not super important, and given the potential for confusion that you have pointed out, we have simply removed the reference to jet depth.

Reviewer wrote: *Line 223 and Table 2: Please add the significance level information (e.g. what is >95% significance level?). The correlations are much less meaningful without this information.*

Our response: Thank you for pointing out this omission. We included p-values in Table 1 and confidence intervals in Fig. 4, but for some reason neglected to do so for Table 2. However, instead of listing p-values in Table 2, we have listed the 95th percentile confidence intervals (we include the p-values in the Table 2 caption). We made this choice

for two reasons. First, we have some concerns about whether the p-values are really meaningful for the results in Table 2. For example, the p-value for PC3 at week 6 is <0.05 despite the fact that the correlation is 0.09; that is, we feel that the 0.09 correlation is essentially useless for prediction purposes despite the fact that it is technically "statistically significant". A second reason to include confidence intervals instead is that they at least provide some amount of information about the "usefulness" of the correlation. Indeed, this is the reason that we included confidence intervals for the similar forecast-verification time series correlation calculations shown in Fig. 4.

Nevertheless, for completeness, we did note what the p-values are for the full time series in the Table 2 caption, but we also added what amounts to a note of caution to readers who might over-interpret the "significance" of the small p-values. Specifically, we included additional p-value information where we tried to take into account autocorrelation in the data that might artificially boost the p-values. This additional calculation shows that indeed the PC3 correlations at weeks 5 and 6 are likely of negligible usefulness/statistical significance.

Reviewer wrote: *There are several instances where the references cited only include recent papers and could benefit by being more historical. For examples: lines 201-202, previously cited work by Reed and Shapiro should be included here too; and Line 358, Olsen, et al., JGR, 2013.*

Our response: Good point. We have updated the citations to be more complete where you have mentioned; we also added the Olsen et al. reference on line 43 (newly added text).

Reviewer wrote: *Fig 10 and related discussion: There are multiple regions and 3 transport types considered. Figure 10 and the discussion concentrates on only 4 of these possible combinations. Are the results not statistically significant for the other area/transports that are not discussed? Or is the discussion representative of the other areas too? Some comment on this should be made in the paper.*

Our response: We did check many other combinations, but for the most part, the other combinations are either for regions nearby to what we are already showing (STT_{500} and TME over the central Pacific) or are not statistically significant. We have added a sentence on lines to make a note of this on lines 355-358.

In addition to this new text, it is worth noting that on lines 412-415 we discuss the fact that TME over Alaska for positive phase EOF1 and negative phase EOF2 appears not to be reliably predictable, while on lines 372-375, we discuss the lack of significant forecast skill over the western US.

Reviewer wrote: *Many of the figures lack labels on the color bars and/or the axes. These need to be added.*

Our response: We find this comment a little confusing because (1) all of the figures have color bars and the units are noted in the figure captions, and (2) all of the figures have axes labels, we just chose to only include the axes labels at the bottom of every column and on the far right-hand side (or sometimes left-hand side) of each row. We prefer to keep this figure label convention as we have it because we believe that it helps to reduce the clutter on multi-panel plots (e.g., Figure 3 would have 18 sets of labels instead of 6, which we find to be preferable).

Reviewer wrote: *The color bar placements in Figure 1 should be better aligned to indicate which figures they are applicable to*

Our response: We have altered figure 1 so that each of the three columns has its own color bar.

Reviewer wrote: *What are the contour levels in Figure 2?*

Our response: Thank you for noticing this omission, we have added contour information to the figure caption.

Reviewer wrote: *Fig. 10: It would be greatly helpful if the color coding of the lines were shown in a legend. Also, why are the medians and whiskers not shown for the verifications? Please include them since it would be useful in the comparisons.*

Our response: Done.

Reviewer wrote: *Lines 138-142: Please add: "...jet variability [in Section 3.1], we compare the EOFs. . ." for clarity.*

Our response: Done.

Reviewer wrote: *Line 192: "loading" likely more appropriately "magnitude"*

Our response: Done.