Response to Anonymous Referee #1

We thank Referee #1 for carefully reviewing our manuscript and providing critical and valuable comments. Please find below responses to each point raised by the Referee.

Comment: p. 4. L. 1. *Monthly precipitation (totals or means?) were derived from presumably daily rain gauge data. It is interesting to look at the number of days per month with precipitation, as the statistical sample of precipitation amounts involves a small (hence greater sampling fluctuations and less well defined mean estimates) in regions with few wet days. To get larger samples, one may use seasonally (3 months) or annually aggregated statistics. Precipitation may be regarded as having two types of statistical distributions: for dry and for wet days. The dry-day statistics is trivial (zero), whereas the wet-day distribution is often described with the gamma distribution (of exponential for a simple approximation). The classification of the data into 'dry' and 'wet' makes sense because different physical conditions are present when it rains and when it doesn't.*

Response: We thank the Referee for this excellent suggestion. Indeed, we initially considered this exact construct to evaluate the predictability of the frequency of precipitation events or wet/dry days, but omitted due to extra manuscript length. However, we have determined that it could be added with minimal extension to the manuscript and are pleased to do so as suggested.

First, for clarification, to this point we have used monthly and seasonal total precipitation, not daily averages, even though the original gauge data is at the daily scale, as correctly presumed by the Referee. (It should be noted, however, that for this study daily data was only readily available from the six stations maintained by SPCC).

We have framed this new piece as complementary to the seasonal total prediction approach as opposed to an entirely unique undertaking. For this auxiliary analysis, we identify the number of wet days as our predictand. As in the original analysis, we use principal component spatial correlation mapping, composite mapping, and global wavelet analysis in an attempt to identify potential predictors of number of wet days.

Interestingly, approximately 85% of the variance experienced in the number of wet days for all stations was captured by the first principal component (PC) of the data. Spatial correlation of the first PC with global SST revealed high correlations within the region of the Pacific typically associated with ENSO, suggesting that the ENSO phenomenon not only controls the seasonal volume of precipitation that reaches the region, but also the number of days during which it falls. No additional regions or climate variables (e.g. sea level pressure, geopotential height, etc.) were identified using this method. Further, composite mapping and global wavelet analysis yielded no additional potential predictors for incorporation into the prediction model. Therefore, the wet-day prediction model uses only the OND Niño 3.4 season-ahead ENSO index as a direct predictor to produce a deterministic prediction of number of wet days in any given JFM season.

New text is added throughout the manuscript in reference to this auxiliary analysis. Specifically, the following will be included in Section 7 of the revised manuscript (now retitled as Additional Prediction Modifications, with three subsections: Extended Lead Time; Spatial Disaggregation of Regional Predictions; and Prediction of Wet/Dry Days):

> "Although predictions of JFM seasonal precipitation totals may be useful for a variety of stakeholders throughout the region, some may prefer additional detailed information such as the frequency of precipitation events expected across a given rainy season. The number of wet or dry days and the intensity of precipitation events can have widespread and serious agronomic/phenologic (Robertson et al., 2008) and infiltration/runoff implications (Mandal and Nandi, 2017). Such information may also be informative to condition stochastic weather simulators for a wide range of hydrologic or agricultural models (Robertson et al., 2006). To evaluate seasonal statistics of wet/dry day frequency for southern Peru, six of the 29 stations having daily data are analyzed. Analogous to the seasonal total precipitation prediction modeling approach, spatial correlation mapping, composite mapping, and global wavelet analysis are all utilized to identify potential predictors describing the expected number of wets days across the JFM season.
>
> Interestingly, approximately 85% of the variance in the number of wet days for all six stations was captured by the first principal component of the data. High correlations are observed between this first principal

component and SST within the region typically associated with ENSO. No additional regions or climate variables (e.g. sea level pressure, geopotential height, etc.) were identified using this method. Further, composite mapping and global wavelet analysis yielded no additional potential predictors for incorporation into the prediction model. Therefore, the wet-day prediction model uses only the OND Niño 3.4 season-ahead ENSO index as a direct predictor to produce a deterministic prediction of number of wet days in any given JFM season. Using the same cross-validation method described previously, the number of wet days per season is predicted for each station.

In addition to correlation coefficients between the predicted and observed number of wet days, the average prediction error for above-average and below-average years is reported for each station. Station-specific statistics are listed in Table 4.

Table 4: Correlation values and average absolute errors for predictions of wet days at each station.

Station (with average number of wet days)	Correlation value between prediction and observation	Average absolute error in years with above average number of wet days	Average absolute error in years with below average number of wet days
1 (25 days)	-0.50	10 days	9 days
2 (36 days)	-0.53	11 days	9 days
3 (51 days)	-0.48	11 days	10 days
4 (54 days)	-0.39	12 days	12 days
5 (54 days)	-0.53	8 days	9 days
6 (33 days)	-0.59	10 days	9 days

Correlations between predictions and observations range from -0.39 to -0.59, with the model performing slightly better in predicting wet days in drier years. In general, however, the simple linear model displays an average absolute error ranging between 8 and 12 days.

To consider overall model performance, a hit miss metric quantifies skill in the two previously introduced categories of years – years with above average number of wet days and years with below average number of wet days (Table 9).

Table 9: Hit miss metric for model predictions of years with above and below average numbers of wet days.

		Predicted conditions	
		W	D
Observed conditions	W	**105**	41
	D	43	**103**

Above average number of wet days (W) and below average number of wet days (D)

Overall, the model correctly predicts whether a given season will have an above or below average number of wet days over incorrectly predicting, with an accuracy of ~72%. The model, however, has a notable bias towards over-predicting near normal conditions (Fig. 13).

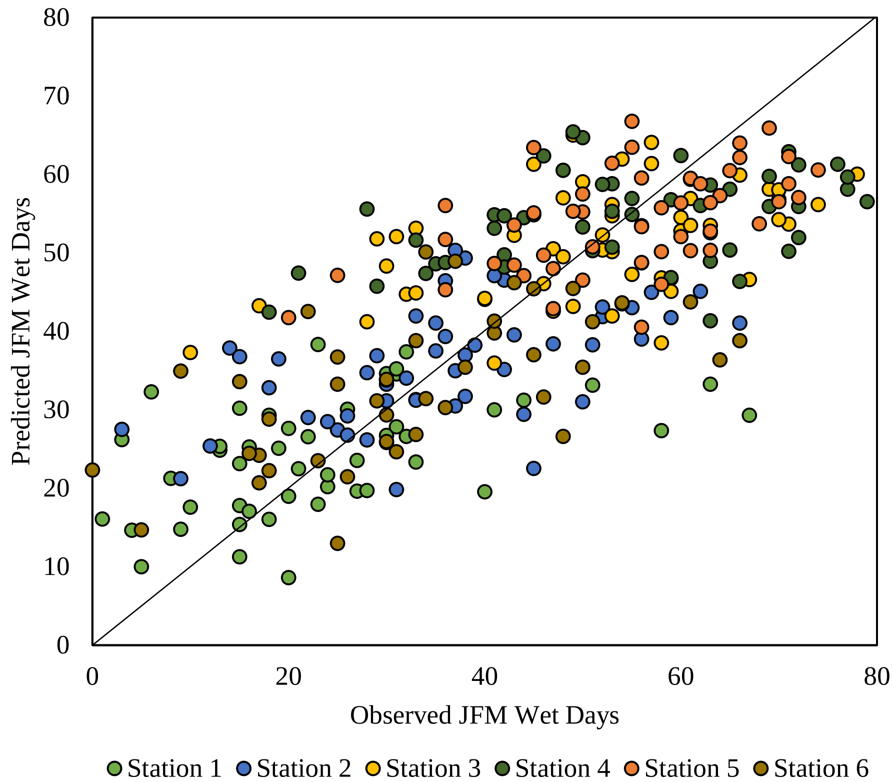

Figure 13: Observed number of wet days compared to predictions.”

In addition, we propose adding the following paragraph to Section 8 (Discussion) of the manuscript:

> “The high correlation observed between number of wet days and SST in the equatorial Pacific suggests that the ENSO phenomenon not only controls the regional seasonal volume of precipitation, but also the frequency of wet days. While the prediction model of wet days achieves notable skill, the model in general displays a tendency to under-predict the number of wet days (especially in seasons in which more wet days were observed than dry). More benefit may be gained in using this model at locations with more wet days (such as stations 3 and 4) as opposed to drier stations. Additional prediction skill may be achieved by incorporating local variables such as antecedent soil moisture conditions or low level winds into station-specific models, as is the case with the disaggregation of regional precipitation predictions.”

In addition to precipitation event frequency, the Referee also comments on precipitation distribution for wet days. Indeed, this is legitimate; however, we have opted not to include this additional analysis here as we are presently less focused on understanding daily precipitation statistics or fitting stochastic weather generators. That said, it does lay the groundwork for others that may be focused on such applications.

Comment: p. 4. L. 6-11. *Please specify if it is version 1 of the reanalysis. Also, the time period covered and the area selected are important. It is important to have sufficient information so that the analysis can be replicated by others independently. Some of this is discussed further down, but it may be easier for the reader if this is provided in a methods section before the results.*

Response: We acknowledge that we did not initially make this clear and have thus updated the manuscript to explicitly state that version 1 of the reanalysis data is used in this study. Additionally, we have clearly noted the period covered by this dataset (1948 to present). To facilitate introducing this information earlier, as suggested, the geographic areas selected and listed in Table 1 of the paper are now referenced in Section 2 (Data Descrpition).

Comment: p. 5. L. 8. *How much of the variance do the subsequent modes capture, and presumably the second order suggests a bi-pole type pattern? There is no need to show this, but perhaps worth describing its character. It is interesting that the leading PC so closely reproduces the station (not area?) mean precipitation. What does that suggest? That the precipitation is dominated by large-scale climatic phenomena (at least aggregated over 3 months) and that other modes are essentially regional perturbations from the large-scale precipitation? I think it may be worth commenting these aspects, but perhaps later in the discussion.*

Response: The Referee highlights a good point, and we agree that it should be further clarified and discussed. The first mode clearly explains the majority of the variance in data (50%), and the second mode captures an additional ~20% of variance; however, the third drops to ~5%. Only these three modes are investigated in this study, for a cumulative total of 75% of variance captured (Fig. 1). The manuscript has been revised to state the variance explained by each mode considered. Indeed, as suggested by the Referee, the second PC suggests a dipole pattern; this description has also been added to the revised manuscript.

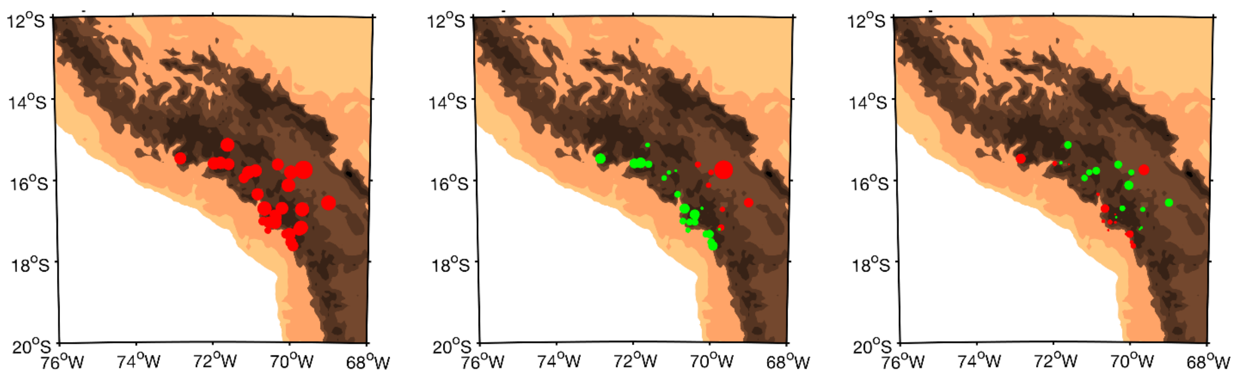

Figure 1: From left to right, spatial pattern produced by the first, second, and third modes of the PCA performed on regional precipitation, explaining ~50%, ~20%, and ~5% of the variance, respectively.

As referenced in the manuscript, Eklundh and Pilesjö (1990) postulated that high correlations between the first EOF of gridded precipitation and area averaged precipitation may suggest the presence of a large-scale climatic phenomena acting homogenously on regional precipitation. Additional studies that support this notion include Ogallo (1980), Mallants and Feyen (1990), Bisetegne et al. (1986). While we use a station average precipitation time series (and not area averaged, as noted by the Referee), the high correlation coefficient between this time series and the first PC of the original set of data may still be interpreted as a widespread homogenous influence on regional precipitation by a large-scale climatic phenomenon. With a correlation between the first principal component of regional JFM precipitation and JFM Niño 3.4 of -0.52, it is likely that this PC describes the modulation stemming from ENSO. As the Referee mentions, the subsequent PCs likely describe regional and local perturbations.

The additional studies mentioned in this response are now referenced in the manuscript.

Additionally, we will add text to Section 8 (Discussion) interpreting the character of each of the PCs considered.

Comment: p. 6. L. 10. *Perhaps the higher modes of the PCA shows the orographic effects.*

Response: The Referee is right that the orographic effects may be evident in various spatial patterns (EOFs), and particularly likely in the higher order modes given that they are more likely associated with regional and local influences. Although this characteristic is not fully explored here, we have added a sentence to the manuscript in Section 3 (Southern Peru Rainy Season and Large-scale Climate Influences) acknowledging that this may indeed be the case:

> "While the first PC likely illustrates ENSO's influence on regional precipitation, it is possible that higher order modes may describe other climatic and topographic forcings such as large-scale climatic phenomena or observed orographic effects."

Comment: p. 8. *Fig. 6: the scatter suggests that more than one factor affects the precipitation, but there is also a discernible anti-correlation between Nino3.4 and the precipitation. An ordinary linear regression can quantify the relationship (and associate a p-value), as can the correlation coefficient. This can be repeated with a subset of the data where the three outliers are excluded to estimate how exceptional they were.*

Response: The Referee is right in assessing that Nino3.4 is highly influential but also not the only relevant signal in descripting precipitation variability. The correlation between JFM precipitation and JFM Niño 3.4 (i.e., concurrent conditions) is -0.57 (p-value = 0.000013). Due to the autoregressive nature of SST in general, Niño 3.4 maintains a relatively high correlation with JFM precipitation for the prior season as well. This has been added to the text and Fig. 6. If the three outlier years (1973, 1990, and 2014) are removed, as suggested by the Referee, this produces a stronger correlation coefficient of -0.66 (p-value = 0.000009), supporting the two-fold notion of ENSO's influence on seasonal precipitation as well as the presence of additional climatic factors that modulate the region's precipitation. This evaluation has also been added to the text immediately following to estimate the importance/influence of these three highly anomalous years, and further underscores the importance of creating a model conditioned on several potential predictors rather than just an ENSO index.

Comment: p. 11. L. 11-16. *Subtracting 9 driest season during El Niño years from the 9 wettest from La Niña years is bound to produce an ENSO signal by design of the analysis. This paragraph seems to be repeated on p. 12. L. 1-5.*

Response: The Referee highlights a valid point. The example composite map, as framed in the manuscript, produces an ENSO signal because only El Niño years and La Niña years are selected for compositing. Thus, this example only marginally supports the points emphasized in this section. Thus, we have removed this figure and text and substituted instead with analysis demonstrating the presence of ENSO through compositing simply the driest and wettest years, regardless of ENSO phase and strength. Accordingly, the manuscript has been revised as follows, including a new Figure 8:

> "Composite maps illustrate climate conditions for a single period or subset of periods, and may be especially useful for understanding forcing mechanisms in anomalous periods. For example, OND SST for the nine subsequent driest JFM seasons on record for southern Peru subtracted from OND SST for the nine subsequent wettest JFM seasons on record for southern Peru produce large positive anomalies in the equatorial Pacific Ocean. This composite map (Fig. 8) further indicates the potential importance of ENSO in explaining JFM precipitation variability in the study region.

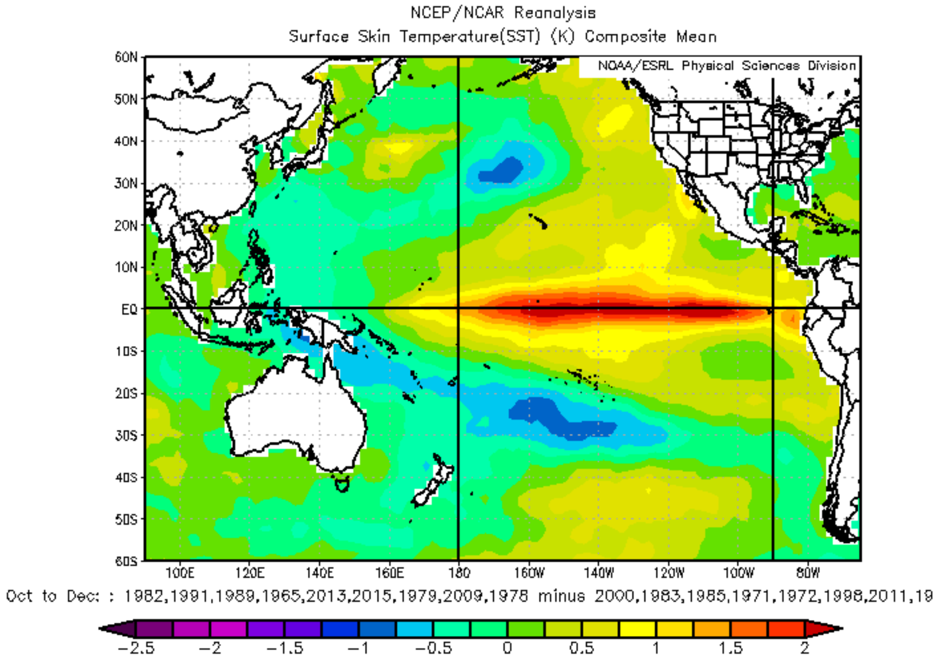

Figure 8: Composite SST OND conditions of nine subsequent driest JFM seasons (1983, 1992, 1990, 1966, 2014, 2016, 1980, 2010, and 1979) and subtracted SST OND conditions of nine subsequent wettest JFM seasons (2001, 1984, 1986, 1972, 1973, 1999, 2012, 1974, and 1997)."

Additionally, any unintended repetition with regard to composite mapping has been removed from the manuscript.

Comment: p. 13. L. 10-14. *It is evident that the PCA applied to area means according to Table 1, but this should also be stated more clearly in the main text. Also, the text should explain how different units are handled - was the PCA applied to standardized indices?*

Response: Thanks for pointing out the need for better clarity. The text has been revised to include reference to the fact that PCA is applied to area mean values of the selected potential predictors. Also, all predictor values are standardized in the PCA process so as not to cause any unintended preference. The text has been appropriately revised.

Comment: p. 14. L. 16-17. *It is not clear how hindcasts were generated from residuals from cross-validation. Please elaborate.*

Response: We agree that this may not have come across clearly in the original manuscript. Hindcasts for precipitation prediction are performed using principal component regression in a drop-one cross-validated mode. This includes – for each year of the hindcast – dropping the predictor data (Z) from the year being hindcasted, forming new PCs (and EOFs) conditioned on the remaining years, and fit to observations using multiple regression, providing an intercept coefficient, regression coefficients, and error term. The predictor data (Z) from the year dropped are then projected onto the EOFs to provide PCs for the dropped year. Finally, these PCs are multiplied by the appropriate regression coefficients and added to the intercept coefficient to provide a deterministic precipitation prediction for the dropped year. This is repeated for each year.

To create ensemble hindcasts, error terms from all years are assembled and a distribution is fit (using a kernel density estimator; the distribution is approximately Gaussian). For each hindcast year, 1000 random draws from the distribution are added to the deterministic precipitation prediction to form an ensemble.

The manuscript has been revised accordingly to clarify this process.

Comment: p. 19. L. 6-8. *From a sample of test results, one will expect some to score high from pure chance. If there are 100 stations, one would expect 5 to score above the 95% confidence level by chance - this is how the confidence interval is defined.*

Response: The Referee makes a valid point. The manuscript text has been modified appropriately to reflect this, as follows:

> "None of the stations experience statistically significant decreases in correlation as a result of disaggregation. For two of the stations, correlation values between station-level predictions and station-level observations increase (statistically significant at the 95% confidence interval) as compared with correlation values between regional-level predictions and station-level observations. (However, it should be noted that there is a 5% probability that this happens by chance.) The remaining stations exhibit no statistically significant change in correlation as a result of station-level scaling."

Minor Comments (all accepted and corrected in the manuscript)

p. 4. L. 2. *'multi-regression' should perhaps be* <u>*'multiple regression'*</u> *or 'multivariate regression'*

p. 5. L. 1. *EOFs often refer to principal components analysis (PCA) of gridded data, weighted by the gridbox area. PCA is mathematically the same thing, but a term used more generally than EOFs. However, this is a matter of taste.*

p. 5. L. 8. *Unless the precipitation has been gridded to onto a regular mesh, the term 'area average' should be replaced with 'station average'.*

p. 19 L. 12. *Split discussions and conclusions into two sections. The conclusions should be brief and repeat the main findings.*

References

Bisetegne, D., Ogallo, L., and Ininda, J.: Rainfall characteristics in Ethiopia, Technical Conference on Meteorological Research in Eastern and Southern Africa, Nairobi, Kenya, 1986.

Mallants, D., and Feyen, J.: Defining Homogenous Precipitation Regions by Means of Principal Component Analysis, Journal of Applied Meteorology, 29, 892-901, 1990.

Mandal, A., and Nandi, A.: Hydrological Modeling to Estimate Runoff and Infiltration in Southeastern Appalachian Debris Flow Complex., 3rd North American Symposium on Landslides, Roanoke, Virginia, USA, 2017.

Ogallo, L.: Regional classification of East African rainfall stations into homogeneous groups using the method of principal component analysis, as in Ikeda, S. et a1. (eds), Statistical Climatology. Developments in Atmospheric Sciences, 13. 225-266, Elsevier, Amsterdam, 1980.

Robertson, A., Kirshner, S., Smyth, P., Charles, S., and Bates, B.: Subseasonal-to-inderdecadal variability of the Australian monsoon over North Queensland, Quarterly Journal of the Royal Meteorological Society, 132, 519-542, 2006.

Robertson, A., Moron, V., and Swarinoto, Y.: Seasonal predictability of daily rainfall statistics over Indramayu district, Indonesia, International Journal of Climatology, 29, 1449-1462, 2008.