Thank you for taking time to carefully read and evaluate our manuscript. We appreciate the comments and suggestions made by reviewers.

Reviewer #2 suggested considering another journal, but we still think this journal, *Climate of the Past*, is the right place to discuss statistical analysis results of paleoclimate data. There is no journal specializing in statistical paleoclimate analysis, and this study applies standard statistical methods to paleoclimate data, so a purely statistical journal is not appropriate. It is our hope that work such as this one can find a home in *Climate of the Past*, as we support the open-access and non-profit nature of that journal and believe statistical work is an important development in paleoclimatology.

Accordingly we have significantly modified our introduction to emphasize the utility of our approach for paleoclimatic analyses and revised the text to give the non-statistical reader a more intuitive grasp of our interpretations.

Modified Introduction:

Paleoclimatic time series hold the promise to extend our knowledge of oceanic, atmospheric, and ecological variability on the timescale of decades to centuries, a time window poorly constrained by instrumental observations. A frequent assumption of such studies is that significant modes of variability detected in the historical record (ENSO- El Nino/Southern Oscillation, PDO- Pacific Decadal Oscillation, AMO- Atlantic Meridional Oscillation, etc.) persist into the past and there may also exist other fluctuations detectable only through the paleoclimate record, but that resemble modern patterns (Knudsen et al., 2011; Koutavas and Joanides, 2012; Newman et al., 2016). Such patterns often involve an ensemble of couplings between aspects such as pressure gradients, surface temperature, and biological productivity--all of which might be observed by sediment proxy methods--and couplings or correlations might be used to infer that the underlying variability does in fact resemble documented modes of modern internal climate variability.

This paper examines statistical aspects of a long-duration, high-resolution, multi-dimensional time series that record variations in among sea surface temperature (SST), phytoplankton productivity, and intensity of the oxygen minimum zone (OMZ) over the Holocene epoch (0.60 to 9.44 kA B.P.) along the central Peru margin, a region strongly affected by the modern ENSO and PDO cycles (Brink et al., 1983; Newman et al., 2016). The sediment is sampled at high-resolution to amount to roughly 3-year averages sampled every 7 years under the accumulation rate typical of the region, and the different proxies are analyzed using the same core sampling, so that correlations are robust of age model uncertainties. These records indicate both surface and subsurface variability in the physical and biological state. Simple correlation analyses revealed that proxy associations vary over time; in some intervals following expected ENSO-like correlations of low sea surface temperature (SST) and enhanced biological productivity; in other intervals this correlation is not apparent (Chazen et al., 2009). We develop here a statistical framework to analyze the evolving relationships over the Holocene in order to interpret proxy correlations, characteristic timescales of variability (“regimes”), and predictability.

Some of the key questions that may be addressed by time series analysis in this region are whether the variability arises from a local or internal source, such as variation in physics through mixing or eddies at the surface (Brink et al., 1983, Colas et al. 2012) or changes in the biological makeup of ecosystems in the region (e.g., Gooday et al. 2010), or from a remote or external source, such as variations in the water properties arriving at the site through large scale modes such as El Nino or the Pacific Decadal Oscillation (Mantua et al. 1997, Deser et al. 2010). The site (Fig. 1) is known for wind-driven upwelling (Brink et al. 1983) at depths shallower than 250m and low oxygen concentrations at depth typical of the eastern tropical South Atlantic oxygen minimum zone, which has been highly variable near 250m depth in recent times (Stramma et al. 2008). Despite the low oxygen levels at depth, the typical sediment accumulation rate over the Holocene during these samples is high (70 cm/kyr), which suggests high, sustained biological productivity and presumably a persistent level of oxygen demand.

A visual analysis of the proxy records (Fig. 2) suggests that the variability of four proxies might fall into multiple regimes: one state with high variability and another state with low variability. This *biphasic* behavior guided our initial analysis using a Hidden Markov Model (HMM; Rabiner, 1989). Hidden Markov methods are increasingly used as a statistically robust automated method for identifying climate regime shifts (e.g., Majda et al. 2006, Franzke & Woollings 2011, Ahn et al. 2017). A benefit of our approach is that it can objectively identify regimes of paleoclimatic behavior in which correlations between proxies and proxy variance evolve (and perhaps alternate) over time. We also explored the possibility of more than two states, but found that these extra regimes were visited only transiently, so parsimony suggested retaining only two modes.

A less common tool in climate modeling is the autoregressive hidden Markov method (AR-HMM, Hamilton, 1988, 1989, 1994) which allows for some memory in the system through a dependence on previous proxy values as well as correlations in the present proxy value noise. The application of this method here, and the insights gained from this application, are a key breakthrough found in our analysis. Both our HMM and AR-HMM results show that there exist two regimes of variability in proxy space at site MW8708-PC2. Here the AR-HMM technique is used to probe deeper into distinctions between causality and correlation, under the premise that a predictive cause should precede its effect in time. As the HMM method examines only simultaneous-in-time correlations, it is not capable of distinguishing causation from correlation in this way. A surprising result of this study is that our conception of the relationships among these proxies from the HMM analysis changed dramatically when the AR-HMM technique was applied and contrasted to the more standard HMM approach. The AR-HMM shows that both climatic regimes show high auto-correlation and low cross-correlation, thereby indicating that none of the proxies are good predictors of other proxies on interannual to decadal timescales. Thus, a hypothesis of local causality between the variables, such as mixing driving local productivity, is not supported by the AR-HMM analysis. This lack of causality is robust to the biphasic regime shifts as well, although in a different location where regime change is not present, a simpler autoregressive only approach can be used to assess causation versus correlation following a similar approach to the methods used here. The software provided with this paper (https://github.com/seonminahn/ARHMM) can be applied for the analysis of multi proxy data from a core record.