

# Training a Non-Homogenous Hidden Markov Model with PMDI and Temperature to Create Climate Informed Hydrologic Scenarios

Burcu Tezcan

Arizona State University, School of Sustainable Engineering and the Built Environment, 660  
College Ave, Tempe, AZ 85281, United States

**Abstract:** Understanding the nature of climatic impacts along spatial and temporal dimensions is important to the development of timely and spatially relevant mitigation options. However, uncertainty in spatial and temporal hydrological patterns on watershed systems is a major challenge in long term water resources planning. This challenge is particularly stark for large watersheds such as the Colorado River, and regions where interbasin transfers and shared demand nodes link multiple watersheds. Here, we developed a non-homogenous hidden Markov model (NHMM) that generates an ensemble of plausible future regional scenarios for any projected temperature sequence. These ensembles can be helpful for water resources managers, infrastructure planners, and government policymakers, with future infrastructure planning and building of more resilient communities when it comes to dealing with natural disasters. Moreover, these ensembles can be used to generate streamflow ensembles, which, in turn, will be a valuable input to study the impact of climate change on regional hydrology. The study presented here contributes towards developing methodologies for creating future wet and dry scenarios at regional scale for large watersheds and regions which are composed of complex interconnecting system networks and exhibit strong climate variability on a variety of time scales and different regions.

## 1.0 Introduction

Uncertainty in spatio-temporal hydrological patterns is a major challenge in long-term water resources planning to inform both infrastructure investments and rules for allocations and operations. This challenge is particularly stark for large watersheds such as the Colorado River, and regions where interbasin transfers and shared demand nodes link multiple watersheds (National Research Council, 2007). Increased understanding of spatial correlation of hydrological variables can improve water management by assessing the probability of co-occurrence of drought across a regional watershed or multiple interlinked watersheds.

Data limitations challenge quantification of regional spatiotemporal patterns from field measurements (Betterle et al., 2017; Blöschl et al., 2013; Razavi & Coulibaly, 2013; Sivapalan et al., 2003). A variety of statistical methods have been applied to the regionalization of hydrological variables such as regression analysis (Merz & Blöschl, 2004), proximity-based methods (Betterle et al., 2019), geostatistical techniques (e.g. inverse distances, kriging, space-time models) (K. Adamowski & Bocci, 2001; Bourges et al., 2012), and combination of time series and spatial statistical methods (J. Adamowski et al., 2013). Despite the flexibility of statistical methods, they are limited by computational requirements and data availability (Blöschl et al., 2013). Physically-based classification frameworks also have been used to characterize regional hydrological variables (Doulatyari et al., 2017). Critically, these methods are limited by the availability of observational data.

Paleoclimate reconstruction records are a promising approach for supplementing relatively short observational records to better understand the long-term climate variability. These reconstructions have been employed to inform water resources planning and policy especially in the areas which are prone to high levels of spatiotemporal variability or in case of modeling supply systems that span multiple river basins (Carrier et al., 2013; Rice et al., 2009; Woodhouse & Lukas, 2006a, 2006b). Reconstructions of the modified Palmer drought severity index (PMDI) and temperature across North America are examples of such records (Cook et al., 2010; Wahl et al., 2012). Recent studies use Living Blended Drought Atlas (Cook et al., 2010) to reconstruct streamflow across the U.S (Ho et al., 2017, 2018). While most applications of streamflow reconstruction are focused on one site, there is a need to address inter-site dependencies for better water supply management of interconnected networks of watersheds for regional water supply. A variety of methods (e.g. Hierarchical Clustering Method, Hidden Markov Model) have been used to quantify the spatiotemporal structure of paleo reconstructed streamflow (Bracken et al., 2016; Carrier et al., 2013; F. Chen et al., 2019; Ho et al., 2016; Rao et al., 2018). For example, Bracken et al. (Bracken et al., 2016) reconstructed flows at 20 sites in the Upper Colorado River Basin demonstrating the ability to preserve inter-site correlations and dynamic representation of uncertainty in each reconstructed year. Progress in regional paleohydrological reconstructions opens new opportunities to integrate a better understanding of regional hydrological patterns to inform decision making, policy, and management practices for networked river systems. However, given the potential for hydrological change driven by a changing climate, an understanding of future patterns is also needed. Annually resolved temperature reconstructions, like PMDI reconstructions, improve understanding of hydroclimatic variability over time scales. Integrating temperature into spatiotemporal models of hydroclimatic variables also enables generation of plausible future scenarios by using the General Circulation Model (GCM) temperature projections.

Critically, there are difficulties in projecting spatial patterns of hydrological variables across multiple sites under future climate scenarios using the GCMs (Vallam & Qin, 2017), as projections are coarse and unsuitable for regional studies (Fowler et al., 2007; C.-Y. Xu, 1999). To overcome this challenge, Statistical Downscaling methods have been used for climate change impact assessment at smaller spatiotemporal scales. These methods are grouped into regression modeling, weather generators and weather typing schemes (Vrac & Naveau, 2007). Alternatively, Regional Climate Models which are nested in the GCMs can be employed for deriving climatic variables for a specific region using Dynamic Downscaling methods which are computationally expensive and only available for limited regions (Salehnia et al., 2019; Sunyer et al., 2012; Tisseuil et al., 2010; Z. Xu et al., 2019). Dynamic Downscaling methods yield spatially distributed fields of climatic variables while preserving certain spatial correlations and maintaining physically realistic relationships between climatic variables (Maurer & Hidalgo, 2008). However, Regional Climate Models are still too coarse for regional hydrological applications for which local climate scenarios are necessary (J. Chen et al., 2011).

A variety of statistical methods have been developed to address hydrological variables regionalization. A variety of methods such as Hierarchical Clustering Method, Hidden Markov Model (HMM), and wavelet analysis have been employed to quantify spatiotemporal variabilities while considering climatic states and trends. While in the Hierarchical Clustering method a predetermined number of clusters does not need to be specified (Ho et al., 2017), the

HMM method provides more useful information in comparison to methods that would seek to cluster the time series or find lower dimensional patterns (Ho et al., 2018). Moreover, HMM has been widely used for downscaling hydrological variables from GCMs, weather generation and modeling reconstructed paleoclimate data. The use of HMM in hydrological modeling is motivated by the ability to capture the regime-switching behavior which is driven by large-scale climate features (Bracken et al., 2014). To quantify spatio-temporal variabilities and understand climatic trends, state space models, such as HMMs have been used extensively in climate science (Zucchini and Guttorp, 1991; Hughes and Guttorp, 1994; Prairie et. al., 2008; Bracken et. al., 2014, 2016; Holsclaw et. al., 2017; Ho et. al., 2018; Hernandez et. al., 2020). While the state transition probabilities and state distributions are held fixed in classical HMM, the nonhomogeneous HMM (NHMM) allows the probabilities or distributions to evolve in time based on large-scale atmospheric predictor variables (Hughes & Guttorp, 1994). Such predictor variables offer one way to link the HMM, which is constructed based on past observations, to future projections.

Previous applications demonstrate that HMM can simulate regional hydrological patterns in space and time. However, these models are trained on historic data and do not have the capacity to make projections or create scenarios to explore the future. The research questions of the paper are: (1) how to generate future hydrological scenarios across the western U.S consistent with historic spatio-temporal patterns and future temperature changes? (2) how to cope with data and computationally intensive NHMMs, particularly for a large region? In answering these questions, we expand upon prior work by constructing a NHMM that can both characterize historical spatiotemporal patterns and make projections informed by temperature projections. Temperature is selected as a predictor variable because it is both available in the paleoclimatic record (Wahl et. al., 2012) and a well vetted output of GCM projections (Woldemeskel et al., 2016). We also apply principal component analysis (PCA) to address the computational complexity of NHMM. Further, our approach improves upon traditional time series models which are based on short-term memory and stationarity that cause a weak persistence and lower probability of long wet and dry spells due to a weak autocorrelation (Bracken, 2014). We address these research questions in the context of the western United States (U.S). The western U.S. region is an ideal test case because of its uncertain climate variability which affects both water resources planning and management.

The following section describes the study area, outlines data and analysis methodology. We then present our results from k-means clustering analysis, PCA, and NHMM. We close by discussing the results and limitations of the study and summarize our conclusions.

## **2.0 Materials and Methods**

### ***2.1 Study Area***

This study focuses on the western U.S. watersheds, particularly the Colorado River and interconnected watersheds linked by interbasin transfers and shared demand nodes, such as Rio Grande River Basin, Central Valley Water Project, Southern California, Los Angeles Aqueduct source watersheds, Central Utah and Strawberry water project, Arkansas River, South Platte River, Little Snake River, Imperial and Coachella Valleys, and part of Mexico. The interconnected nature of these watersheds motivates this investigation of regional hydroclimatic patterns, as droughts occur at larger spatiotemporal scales for the western part of U.S, typically

spreading over hundreds to thousands of square kilometers. Droughts in the western and most of the central USA is originated from northwestern direction (Konapala and Mishra, 2017). Hence, the study area is defined as 30 – 49° N, 97.1 – 124.9° W (Figure 1).



Figure 1. Map of the Study Area with the HUC-2 hydrologic regions (green polygons). The study area is defined as 30 – 49° N, 97.1 – 124.9° W.

Figure 1 shows the study area along with the 11 U.S. Geological Survey (USGS) HUC-2 (2-digit hydrologic unit code) hydrologic regions. Spatially gridded PMDI and temperature data are extracted based on the HUC-2 regions. More detailed information on the HUC-2 regions is listed in Table 1.

Table 1. Summary of HUC-2 regions for the study area

Code	Name	Area (10 <sup>5</sup> ) km <sup>2</sup>
17	Pacific Northwest	7.3
14	Upper Colorado	2.9
15	Lower Colorado	4.2
13	Rio Grande	3.4
16	Great Basin	3.7
12	Texas-Gulf	2.2
18	California	4.4
10	Missouri	11
11	Arkansas-White-Red	3.8

## 2.2 Data

PMDI reconstructions used in this study are obtained from the National Oceanic and Atmospheric Administration (NOAA) database (Cook, et. al., 2010). Temperature projections are taken from the bias corrected fifth phase of the Coupled Model Intercomparison Project (CMIP5) climate projections with the Representative Concentration Pathway (RCP) 4.5 scenario, which represents an approximate doubling of carbon dioxide levels relative to pre-industrial

levels by 2100 (Voldoire, 2013). Choice of study period is guided by a trade-off between data availability in terms of record length and spatial coverage. Annually resolved paleoclimate records and temperature projections are used to test hydroclimatic variability over time scales related to water resources management and planning. The overview of data used in this study is given in Table 2.

*Table 2. Overview of Data*

<b>Data Type</b>	<b>Data Period</b>	<b>Source</b>
PMDI reconstructions	1500 - 1980	Cook et. al., 2010
Temperature reconstructions	1500 - 1980	Wahl et. al., 2012
Temperature projections (CMIP5 – RCP 4.5)	2020 - 2100	Voldoire, 2013

### *2.2.1 Living Blended Drought Atlas*

The PMDI is a modification of the Palmer Drought Severity Index (PDSI), which uses readily available temperature and precipitation data to calculate relative dryness. The difference between PMDI and PDSI is the transition periods between dry and wet conditions. For the PDSI, a dry/wet index is calculated when the probability that a drought/wet spell is over is 100%, and transition index is assigned when the probability is less than 100%. The PMDI incorporates a weighted average of the wet and dry index by using the probability as the weighting factor. Both the PMDI and PDSI will have the same value during an established drought or wet spell (i.e., when the probability equals to 100%). However, they will have different values during transition periods since PMDI has a more gradual transition from one spell to the. PMDI is the operational version of the PDSI and the best suited index for operational applications (Heddinghaus and Sabol, 1991). The values of PMDI generally range from -6 to +6, where negative values represent dry spells, and positive values are wet spells.

This product is well validated, and versions of the NADA have been used extensively in the study of North American drought variability (NADA; Cook et. al., 1999). The NADA is composed of annually resolved summer (June - August) PDSI reconstructions from a network of tree-ring chronologies estimated on a 286-point 2.5 x 2.5 PDSI grid over most of North America (Cook et. al., 2007). To better characterize the regional drought variability, an updated version of NADA, the Living Blended Drought Atlas version 1 (LBDv1), was introduced. LBDv1 reconstruction includes additional tree ring chronologies and covers North America at a spatial resolution of 0.5 x 0.5. To include droughts in the 21<sup>st</sup> century, The LBDv1 data is updated until 2017. The Living Blended Drought Atlas version 2 (LBDv2) is based on LBDv1 and calculates a PMDI by recalibrating the PDSI (Cook et. al., 2010). In order to access long term drought data and identify the patterns, the PMDI from the LBDv2 is used in various studies (e.g. Burgdorf et. al., 2019; Son et. al., 2021). In this study, the observational and paleoclimate PMDI records from LBDv2 between 1500 and 1980 are used for the analysis since the availability of observational records is short relative to the time scale of hydrological variability.

### *2.2.2 Temperature Reconstructions*

Temperature reconstructions, like PMDI reconstructions, play a significant role in understanding the climate prior to the beginning of the observational records by quantitatively extending the record back in time. A tree ring-based reconstruction of western North America annual surface

temperature anomalies with a 5 x 5 degree grid cell coverage is used in this study for the period of 1500 and 1980. The spatially explicit reconstructions are calculated based on a truncated empirical orthogonal function method (Wahl et. al., 2012), and have been used in multiple studies (e.g., Anchukaitis et. al., 2013; Lehner et. al., 2017). The reconstructed temperature anomalies are used as covariate in the NHMM in order to enable the model to generate plausible future scenarios by using GCM temperature output.

### *2.2.3 Temperature Projections*

The Intergovernmental Panel on Climate Change Fifth Assessment Report (IPCC - AR5) promoted the development of the fifth phase of the Coupled Model Intercomparison Project (CMIP5) to improve understanding of climate and provide estimation on future climate change consequences. Model simulations encompass the period of 1900 - 2100. During the historical period (1900 - 2005), the CMIP5 experiment simulates temperature with forcings driven by greenhouse gasses and aerosols due to human activities on climate, land use change, changes in solar radiation, and natural sources, such as volcanic activities, solar irradiance (Swain & Hayhoe, 2015). The future period of the CMIP5 climate projections covers the period of 2006 - 2100, and it is bias corrected. In this study, we used the bias corrected 1x1 degree grid cell coverage CMIP5 climate projections with the Representative Concentration Pathway (RCP) 4.5 scenario, which represents an approximate doubling of carbon dioxide levels relative to pre-industrial levels by 2100 (Volz, 2013). Annual mean surface temperature anomalies are calculated to be compatible with the historical temperature reconstructions. In order to enable NHMM to generate ensemble of future regional PMDI for any projected temperature sequence, anomalies are estimated from 1900 to 2100 by subtraction of the average temperature values for each grid cell between 1900 and 1980.

## **2.3 Methods**

In this study we develop a NHMM of PMDI in the western U.S. with temperature as a covariate. To reduce the computational demands, we apply Principal Component Analysis.

### *2.3.1 Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) is a multivariate technique that reduces a data set containing many variables to a data set having fewer new variables. These new variables are linear combinations of the original variables, and they have high variance while being uncorrelated with each other. The PCA method can represent much more compact data variations for atmospheric and other geophysical fields, which exhibit many large correlations among the variables (Wilks, 2011). Many studies have used the PCA in climate and hydrologic sciences to describe dominant patterns of observed records (e.g., Bethere et. al., 2017; Balling et al., 2007; Lins, 1997).

The aim of PCA is to explain a majority of the variance in the original data with a smaller number of variable dimensions, thereby reducing complex temporal and spatial climatic data sets to interpretable abstractions (White et. al., 1991). Some researchers have used PCA to examine spatial variability of wet and dry periods (Eder et. al., 1987; Raziei et. al., 2008; Ogunrinde et. al., 2020; Huang et. al., 2022) while others have applied PCA for dimension reduction.

Malmgren and Winter (1999) used PCA with varimax rotation to seasonal data in order to decrease dimensionality of their data in Puerto Rico. Furthermore, in order to avoid overfitting and multicollinearity, PCA is employed to reduce high dimension of drought related variables and indices (Hao et. al., 2018).

In this study, PCA is used to reduce the dimensionality of large PMDI and temperature reconstruction records, increasing interpretability while minimizing information loss. All data is scaled and centered while performing the PCA. The number of principle components (PCs) are determined by calculating cumulative explained variance ratio, which is a function of the number of components. A scree plot is the common approach to depict this ratio, and select number of PCs (Cattel, 1966). It shows the curve for explained variance related to each successive PC. The criterion of scree plot looks for a break (or elbow in the curve) between the components and selects all components just before the line flattens out. Note that as an alternative method for dimensional reduction, we also tested k-means clustering but the results were not satisfactory for this application. The methodology and results for the k-means clustering analysis are presented in Appendix A.

### 2.3.2 *Non-Homogenous Hidden Markov Model (NHMM)*

Hidden Markov models (HMMs) are statistical models that generate a variable sequence from a distribution based on the state of an underlying and unobserved Markov process (Zucchini et. al., 2009). In the classical homogeneous HMMs, a system changes between unobserved or hidden states characterized by transition probabilities and a Markov chain. Each state corresponds to a probability distribution where observed time series are drawn (Bracken, 2014).

The time-homogeneity of the classical homogenous HMM can be limiting in practice if observations are non-stationary or have seasonal dependence. One approach to relaxing this assumption is making transition probabilities to be dependent on covariate time series, which is called a non-homogenous HMM (NHMM) (Hughes and Guttorp, 1994; Robertson et. al., 2003; Bracken, 2016; Holsclaw et. al., 2017). Temporal inhomogeneity can also be introduced to the emission component of the model by allowing the parameters of the emission distributions to vary with time and location as a function of covariates (Holsclaw et. al., 2017).

According to Hughes and Guttorp (1994), it is possible to describe the historical relationships between atmospheric circulation and a given process using NHMMs. It also allows simulation of space-time realizations of a regional hydrologic process conditional on a sequence of atmospheric data. Analogously, we here create space-time realizations of PMDI conditional on temperature, which is physically linked to PMDI through atmospheric water demand. Furthermore, this enables use of GCM results from climate scenarios to project the impacts of such climate changes on regional hydrological processes. Here, we adopt a Gaussian NHMM, which uses the spatio-temporal historical data (e.g., PMDI reconstructions) and an exogenous variable (e.g., temperature reconstructions). Once the model is fit, it can be applied to create an ensemble of space-time realizations of PMDI conditional on projected temperature changes by linking transition probabilities between states and emission distribution parameters to GCM generated temperature. An ensemble of state sequences is generated based on the transition probabilities by iterating the fitted NHMM 100 times (Figure 2). Then, a unique mean and standard deviation are calculated for each location (e.g., clusters/PCs) for a given state based on the projected temperature at that location and time. Figure 3 shows the structure of the model.

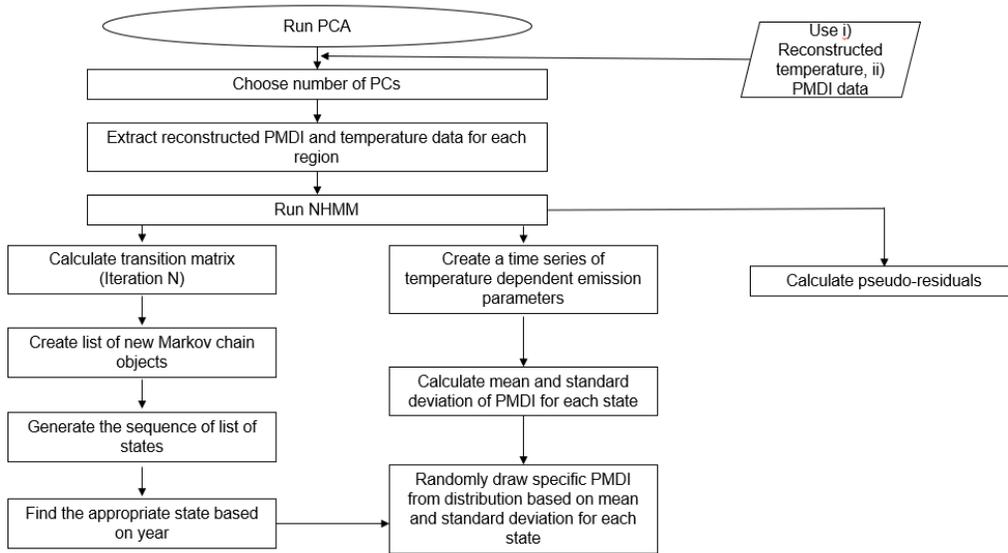


Figure 2. Flow diagram of methodology to fit NHMM and to apply NHMM to simulate PMDI and calculate pseudo-residuals.

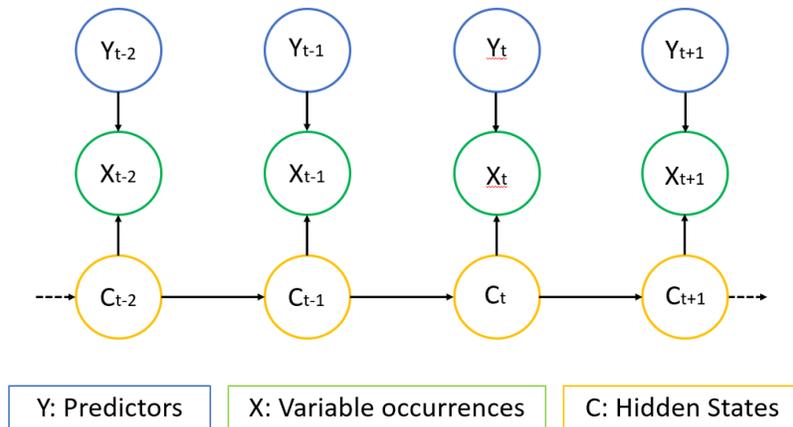


Figure 3. Structure of NHMM.  $Y$  represents the predictors (additional variable, such as temperature),  $X$  represents the variable occurrences (such as, PMDI), and  $C$  represents the number of hidden states (adapted from Zucchini et. al., 2009).

The model is fitted with a range of hidden states, then Bayesian Information Criterion (BIC) and Akaike's Information Criterion (AIC) are used for model selection. BIC seeks to maximize model consistency while AIC seeks to maximize model efficiency (Celeux and Durand, 2008). There are theoretical questions about the use of BIC and AIC in this context. While BIC has a tendency to underestimate the number of hidden states, AIC demonstrates a tendency to overfit the number of hidden states in an HMM (Celeux and Durand, 2008; Buckby et. al., 2020). The appropriate number of hidden states in an HMM can be determined by the minimum BIC and/or AIC value (Bacci, et. al., 2014). However, both AIC and BIC can be used to determine a range of plausible model sizes by model averaging (Diziak et. al, 2020). Here, we select a mid-point of hidden states of models to balance the goals of efficiency and consistency, if the two metrics disagree.

To assess whether the fitted model describes the data well, pseudo-residuals (also known as quantile residuals) are calculated as an additional check on model performance based on the information provided following the procedure detailed by Zucchini and MacDonald (2009). This cannot be done by analyzing only standard residuals because the observations are explained by different distributions depending on the active hidden state. Following Zucchini and MacDonald (2009), we concluded that the observations are modeled well if pseudo-residuals are close to standard normal distribution. We visually assessed the residuals and pseudo-residuals using histograms and the Shapiro-Wilk normality test (Shapiro and Chen, 1968).

### 3.0 Results

The results of NHMM with PCA are presented in this section. The identified optimal number of PCs for the western U.S. are used as an input to NHMM models. The number of hidden states for the NHMM is determined based on the model selection criteria mentioned above. After model fit, pseudo-residual analysis and Shapiro-Wilk test are conducted. The results show that pseudo-residuals are close to normal distribution. After developing the model for the historic training period, multiple plausible sequences of states are generated for the forecast period. Future PMDI scenarios, informed by a specific temperature scenario, are created over the western U.S.

#### 3.1 PCA

The scree plot and the cumulative variance plot criterion are used to select the appropriate number of PCs (Cangelosi and Goriely, 2007) (Figure 6).

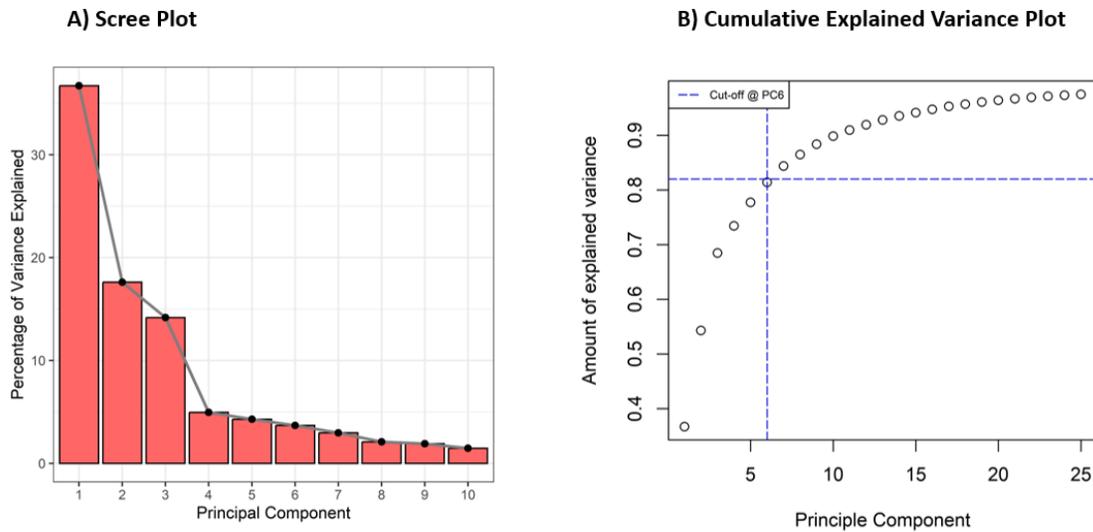


Figure 6. PCA selection metrics for the PMDI reconstructions: A) Scree plot and B) Cumulative explained variance plot.

Six PCs are retained for the PMDI reconstructions, explaining 81.4 percent of the variance. For temperature reconstructions, ten PCs, explaining 80.5 percent of the variance are used. Figure 7 shows spatial patterns of the first six PCs that are used in the NHMM for the PMDI reconstructions.

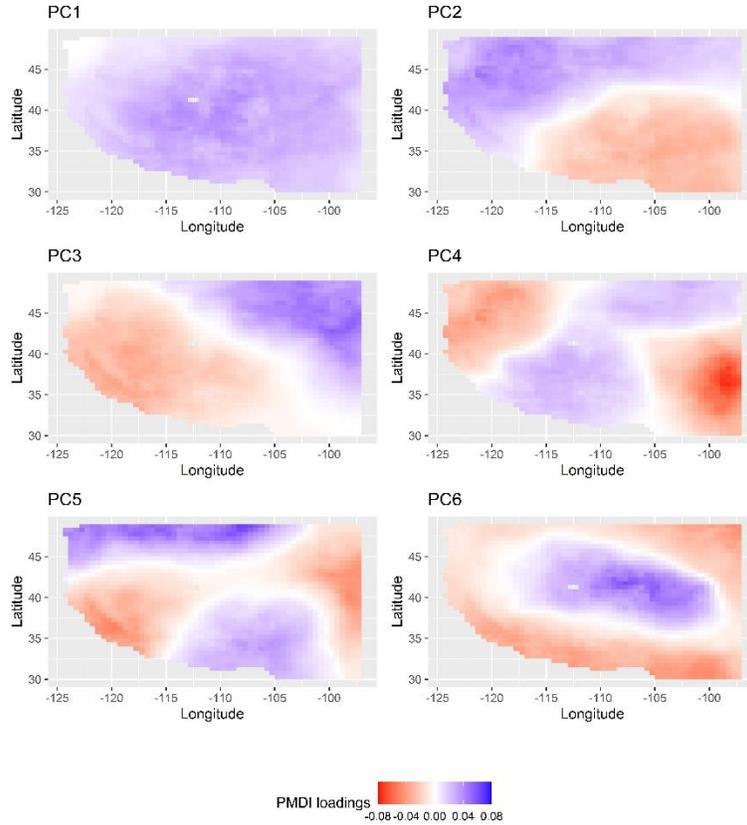


Figure 7. Spatial pattern of loadings of first six PCs based on PMDI reconstructions.

### 3.2 Non-Homogenous Hidden Markov Model

The minimum BIC is calculated at state one and the minimum AIC is found at state eight. A NHMM model with four hidden states is selected as the mid-point between AIC and BIC criteria. (Table 3).

Table 3. NHMM model selection

NHMM State	AIC	BIC
State=1	26945.38	27246.04
State=2	26934.16	27548.01
State=3	26941.9	27877.29
State=4	<b>26955.31</b>	<b>28220.6</b>
State=5	26956.99	28560.52
State=6	26986.19	28936.32
State=7	27025.55	29330.63
State=8	26917.68	29586.06
State=9	27113.9	30153.93

The most likely sequence of states through time is revealed by the Viterbi algorithm. The resulting hidden states enable examination of spatio-temporal patterns over the western U.S. Figure 8 shows PMDI and temperature reconstruction anomaly time series together with most likely hidden states over the western U.S. during training period. Positive PMDI represents wet conditions while negative PMDI represents dry conditions. States represent wet, average, and dry conditions.

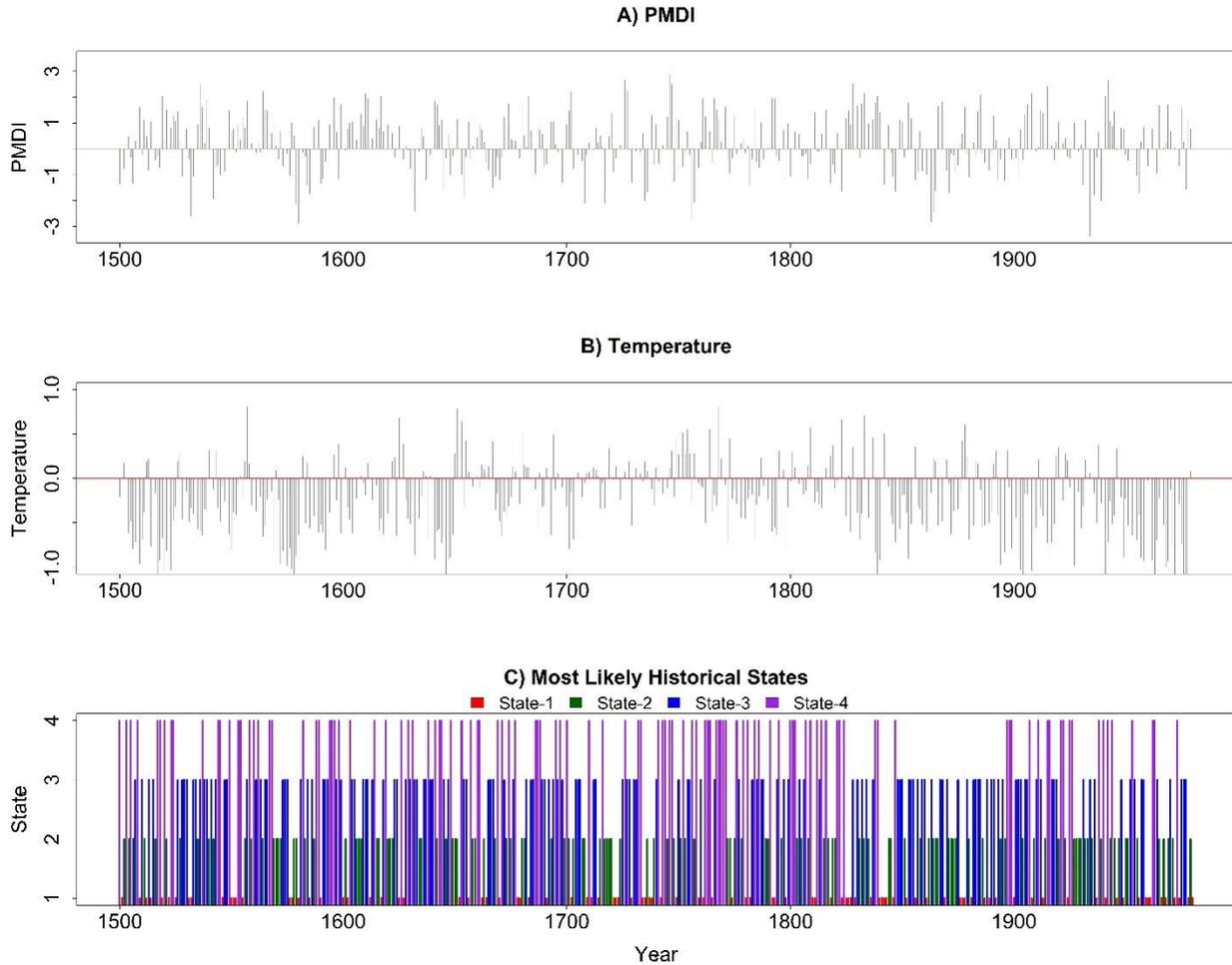


Figure 8. A) Annual mean PMDI reconstruction time series over the western U.S., B) Annual mean temperature reconstruction anomaly time series over the western U.S., and C) Most likely states over the western U.S.

The histograms of the pseudo-residuals are plotted to check whether the pseudo-residuals are normally distributed. All histograms are found to be close to normal distribution due to their approximate symmetric bell-shape. Figure 9 illustrates the distribution of pseudo-residuals for the first PC of NHMM over the western U.S. In addition to histograms, the Shapiro-Wilk test is applied and p-values for pseudo-residuals are found to be greater than the chosen alpha level of 0.05, which confirms the normality of the pseudo-residuals.

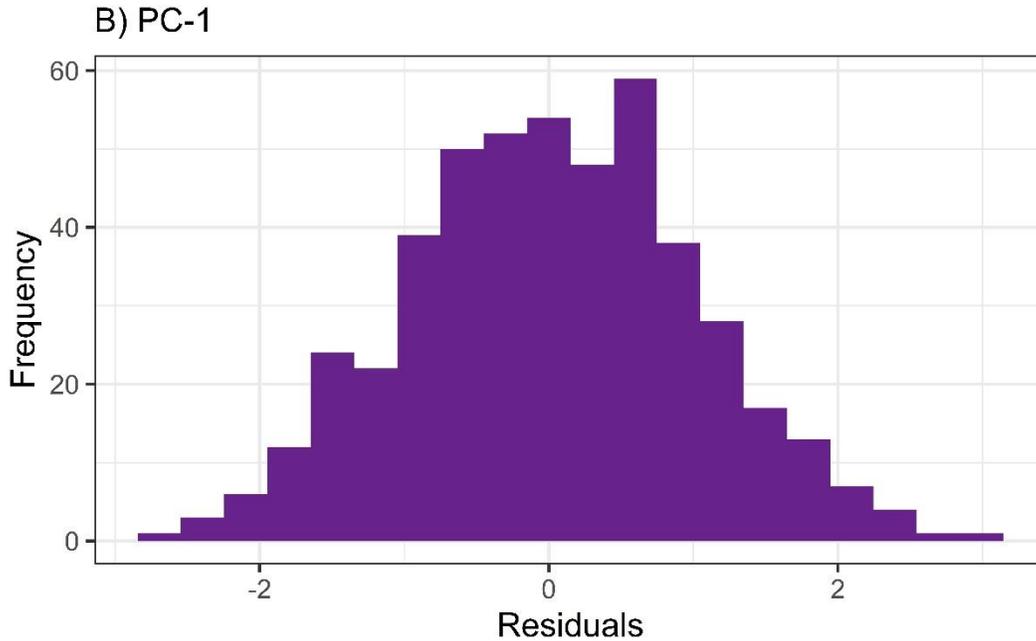


Figure 9. Distribution of pseudo-residuals for the first PC. Distributions of pseudo-residuals for all PCs are close to normal distribution.

Given the model developed for the historic training period, one can generate plausible sequences of hidden states for the forecast period. Figure 10 illustrates the one stochastic scenario of system states (e.g., iteration 1) at each time step during the forecast period. After identifying the system states, the NHMM is used to simulate values of PMDI by sampling from the state distributions. Future PMDI scenarios, informed by a specific temperature scenario, are created from 2020 to 2100 for the western U.S by using the model.

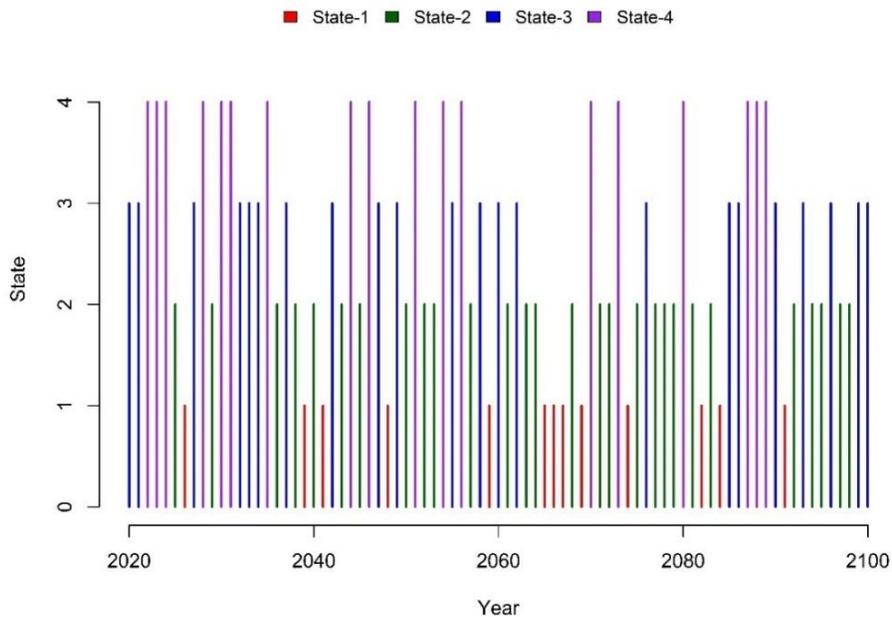
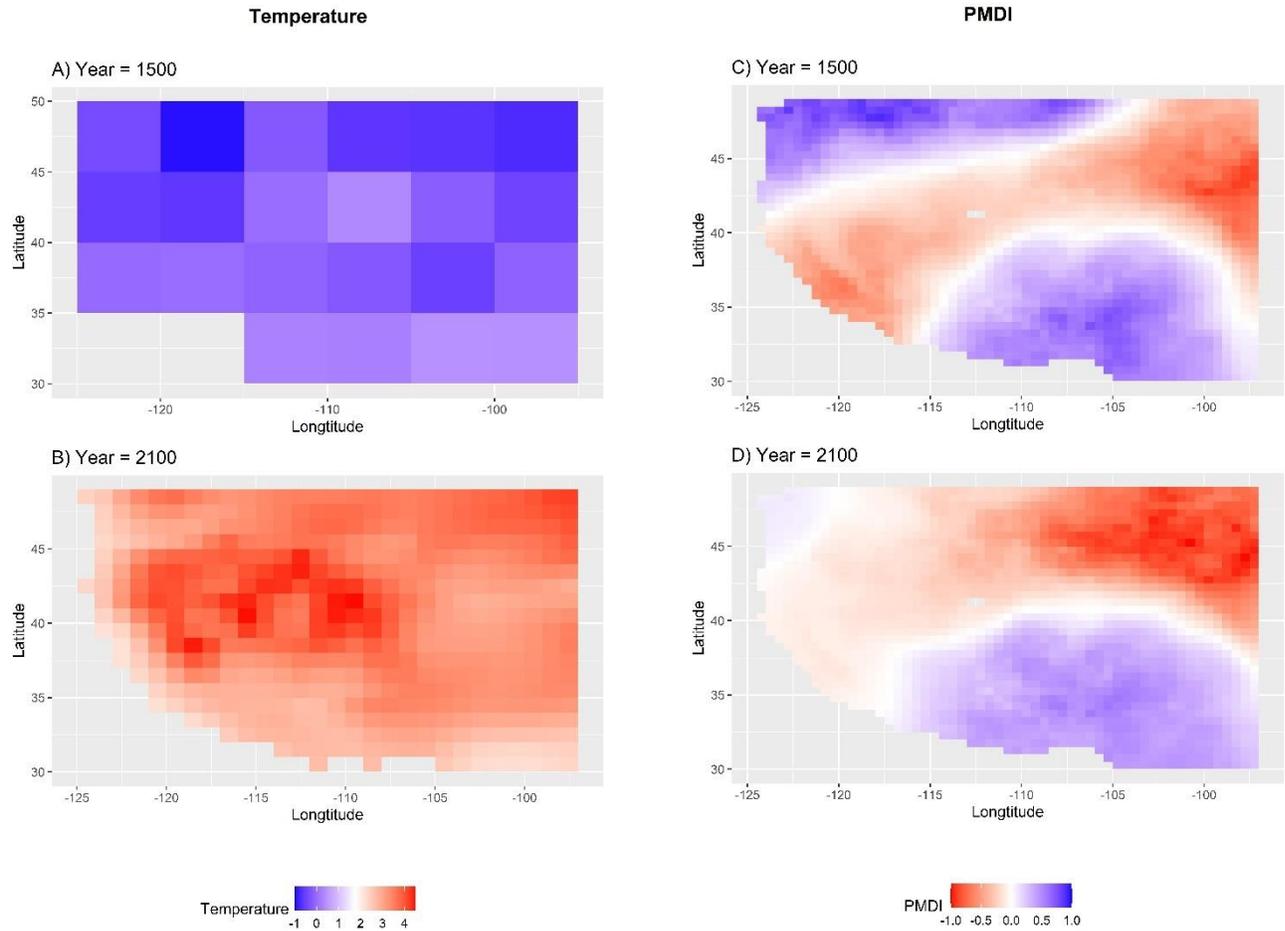


Figure 10. One stochastic scenario of system states in the western U.S. (Scenario – 1).

Figure 11 shows the spatial pattern of annual temperature reconstruction anomalies and simulated annual PMDI reconstructions in the study area for one historic and one projected year. The annual temperature anomaly significantly rises by the end of the century all over the western U.S. under the RCP 8.5 scenario.



*Figure 11. Spatial comparison of annual temperature reconstruction anomalies and simulated annual PMDI reconstructions for years 1500 and 2100 in the western U.S. (Scenario -I).*

Quantile range for ensemble of ten sequences of scenarios over time for the case study area are shown in Figure 12. The figure shows changes in PMDI variability (quantile differences) of each ten scenarios over time. The wide range indicates high variability, and the small range specifies low variability.

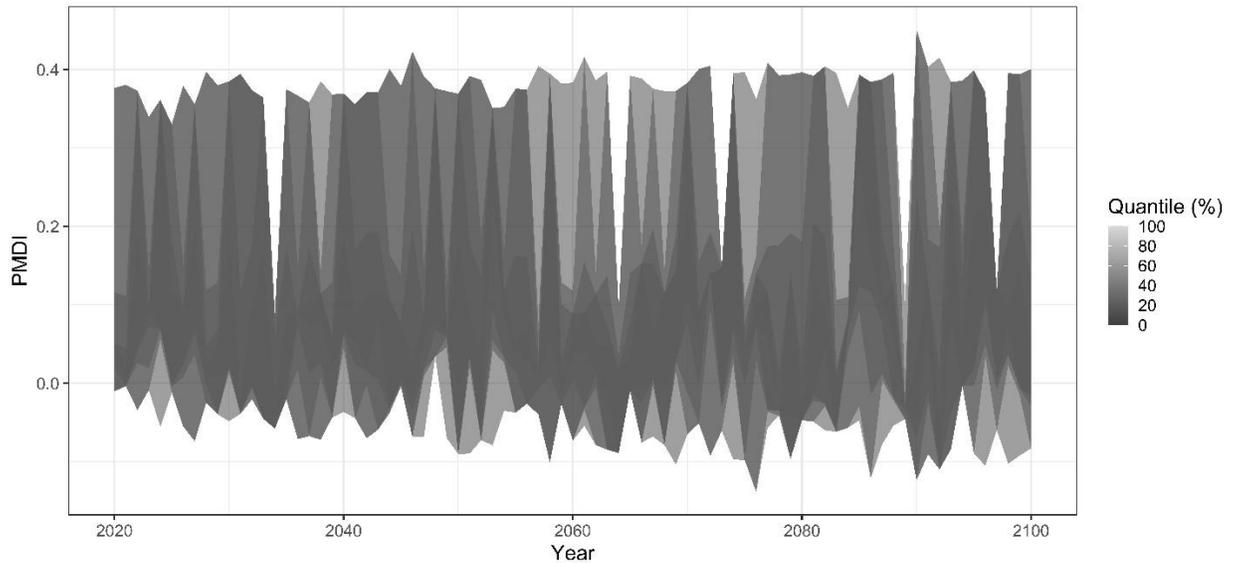


Figure 12. *Quantile range for ensemble of ten sequences of scenarios over time for the western U.S.*

#### 4.0 Discussion

Traditional stochastic time series models are not able to capture regime-switching behavior due to their short-term memory and stationarity. This leads to a weak persistence and lower probability of long wet and dry periods due to a weak autocorrelation (Bracken et al., 2014). Hence, these models misrepresent the risk of prolonged wet and dry periods, consequently effects both water resources planning and management (Bracken et al., 2014). Additionally, some time series models leveraged currently, such as HMMs, perform satisfactorily in capturing the regime-switching behavior (Fortin et al., 2004; Bracken et al., 2014, 2016; Ho et al., 2018), however they fail to reproduce the variability of non-stationary given that exogenous climate predictors are not considered. Using climate-informed variables (e.g., temperature) as covariates can help capture the variations in a hydrological variable (e.g., PMDI) that are influenced by the covariates through physical processes. Thus, utilizing a covariate that shows changes in climate makes potentially more skillful models in capturing persistence and model nonstationary (Ho et al., 2018; Vinnarasi and Dhanya, 2022). One of the contributions of this study is to address the limitation of capturing nonstationary features while modelling the regime switching spatiotemporal behavior. The developed model can capture nonstationary by allowing the parameters of the emission distributions to vary with time and location as a function of a covariate.

Data limitations challenge quantification of regional spatiotemporal patterns (Betterle et al., 2017; Blöschl et al., 2013; Razavi & Coulibaly, 2013; Sivapalan et al., 2003). Having relatively short instrumental records is a key limitation to assess spatial patterns in wet and dry periods in the U.S. for long time scales. Paleoclimate reconstructions offer a long-term perspective on climatic variability by supplementing relatively short observation data (Ho et al., 2018). Annually resolved paleoclimate records provide a framework for exploring policy and

management alternatives to mitigate or adapt the future changes (USGS, 2022). Another advantage of this study is that we constructed the NHMM with paleoclimate time series of PMDI, and temperature records leveraged together with observed data to better understand the long-term climate variability which is relevant to water resources management and planning.

One benefit of using PMDI reconstructions to define wet and drought events is that this index is a modified version of the PDSI, which identifies climate events solely based on hydroclimatic variability, rather than including human impact. The PDSI and PMDI will result in equal values in periods that are clearly wet or dry, but the PMDI will yield smoother transitions between wet and dry periods (Heddinghaus and Sabol, 1991). Thus, PMDI has been used in many previous studies to identify wet and dry periods (Diffenbaugh et. al., 2015; Pongracz et. al., 1999; Wahl et. al., 2022).

Given the high dimensionality of the paleoreconstruction data, PCA is applied to reduce the dimensionality of data and thereby enable computationally efficient NHMM. Precisely, the computation time is seven minutes for the model to generate an ensemble of 100 sequences of scenarios. All experiments are run on a 64-bit computer with Intel(R) Core (TM) i7-8665U CPU @ 1.90GHz 2.11 GHz processor and 32 GB of RAM running Windows 10 Enterprise. Another advantage of developing the NHMM with PCA output is its grid-based representation, which provides insight into wet and dry events by representing a specific PMDI value at a given time and state. As shown in Figure 11, the ability to have grid-based representation is of interest for local and regional resource managers since it generates realistic and spatially variable scenarios.

Accurate simulation of droughts in GCMs are limited due to the chaotic nature of hydroclimatic variables and the complexity of hydroclimatic processes (Ho et. al., 2018). GCMs and regional climate models, which use computationally expensive Dynamic Downscaling methods to derive specific variables from GCMs, are too coarse for regional hydrological applications (Fowler et al., 2007; C.-Y. Xu, 1999; J. Chen et al., 2011). The NHMM simplifies the temporal and spatial structures to be parameterized despite its large number of parameters and computational complexities (Mehrotra and Sharma, 2005). As seen in the Figure 12, the study presented here is able to avoid these issues, creating future PMDI scenarios from year to year at each grid location by linking the model with a predictor variable (e.g., temperature projections) and using computationally efficient model as mentioned above paragraph.

According to GCM outputs and detailed regional studies, the regional impacts of global warming on future water supplies are uncertain and streamflow is sensitive to changes in temperature and precipitation (Frederick and Gleick, 1999). For example, McCabe and Wolock (2007) showed that 1°C to 2°C increases in temperature could result in substantial water supply shortages in the Upper Colorado River Basin. They also reported that future warming could increase the tendency of failure to meet the water allocation requirements of the Colorado River Compact. Creation of ensembles of regional climate scenarios is necessary for the quantification of climate uncertainty in the influence of global warming to address the potential impacts of climate change and climate variability (Groves et. al., 2008), hereby informs future infrastructure planning and water policy. One of the significant contributions of the paper is that the developed model generates an ensemble of plausible future regional PMDI scenarios for any projected temperature sequence. These ensembles can be helpful for water resources managers, infrastructure planners, and

government policymakers, with future infrastructure planning and building of more resilient communities when it comes to dealing with natural disasters. Moreover, streamflow ensembles that preserve long-term spatio-temporal variability can be generated by using these ensembles. Thereby, play a pivotal role to address creating intersite streamflow reconstructions for better water supply management of interconnected networks of watersheds.

An important limitation of this study is that PMDI and temperature reconstructions are based on tree-ring chronologies, and the uncertainty increases with the age of the chronology. Therefore, the variability of the reconstruction data and subsequent state together with emissions distribution is likely to be affected by the changing number of chronologies. To assess whether the fitted model describes the data well, pseudo-residuals are calculated since it allows a comprehensive residuals analysis in Markov-switching models (Zucchini and MacDonald, 2009). The distribution of pseudo-residuals for the first PC of NHMM is presented in Figure 9. Overall, the developed model appears to fit the data adequately due to their approximately normally distributed pseudo-residuals.

## **5.0 Conclusion**

The work reported in this paper contributes towards developing methodologies for creating future wet and dry (or PMDI) scenarios at local or regional scale for large watersheds and regions which are composed of complex interconnecting system networks and exhibit strong climate variability on a variety of time scales and different regions. As a first step, PCA is performed to the grided PMDI and temperature reconstruction data in order to reduce the dimensionality. The model is trained with principal components of PMDI reconstructions and principal components of temperature reconstructions as predictor for a period of 481 years from 1500 and 1980, then most likely hidden states are derived for the western U.S. The developed model is capable of representing adequately the reconstruction data since its pseudo-residuals are normally distributed. The trained model is applied using the projected temperature from GCM output as a forcing variable for a period of 81 years from 2020 to 2100 in order to generate plausible future PMDI scenarios and sequences of states. The developed model effectively models the regime switching spatiotemporal behavior while capturing the nonstationary in the western U.S. Additionally, annually resolved paleoclimate records leveraged together with observed data allow for the examination of long-term climatic variability over time scales relevant to water resources management and planning to mitigate or adapt to future changes. Another advantage of the model is that the NHMM presented here can generate an ensemble of future regional PMDI for any projected temperature sequence. These ensembles can be used to generate streamflow ensembles and to inform water supply planning. The methodology developed here could also be applied using other drought metrics, depending on the application of interest. This approach can be used for any region or watershed to better understand the spatio-temporal patterns of drought events. An ensemble of plausible future regional PMDI scenarios can be used to inform watershed or regional planning and decision making. Furthermore, these ensembles can be used to generate streamflow ensembles, which, in turn, will be a valuable input to study the impact of climate change on regional hydrology and water management.

## References

- Adamowski, J., Adamowski, K., & Prokoph, A. (2013). Quantifying the spatial temporal variability of annual streamflow and meteorological changes in eastern Ontario and southwestern Quebec using wavelet analysis and GIS. *Journal of Hydrology*, 499, 27–40. <https://doi.org/10.1016/J.JHYDROL.2013.06.029>
- Adamowski, K., & Bocci, C. (2001). Geostatistical regional trend detection in river flow data. *Hydrological Processes*, 15(18), 3331–3341. <https://doi.org/10.1002/hyp.1045>
- Ahmed, M., Anchukaitis, K. J., Asrat, A., Borgaonkar, H. P., Braida, M., Buckley, B. M., ... & Zorita, E. (2013). Continental-scale temperature variability during the past two millennia. *Nat. Geosci*, 6(5), 339–346.
- Bacci, S., Pandolfi, S., & Pennoni, F. (2014). A comparison of some criteria for states selection in the latent Markov model for longitudinal data. *Advances in Data Analysis and Classification*, 8, 125–145.
- Balling, R. C., & Goodrich, G. B. (2007). Analysis of drought determinants for the Colorado River Basin. *Climatic Change*, 82(1), 179–194.
- Betterle, A., Radny, D., Schirmer, M., & Botter, G. (2017). What Do They Have in Common? Drivers of Streamflow Spatial Correlation and Prediction of Flow Regimes in Ungauged Locations. *Water Resources Research*, 53(12), 10354–10373. <https://doi.org/10.1002/2017WR021144>
- Bethere, L., Sennikovs, J., & Bethers, U. (2017). Climate indices for the Baltic states from principal component analysis. *Earth System Dynamics*, 8(4), 951–962.
- Betterle, A., Schirmer, M., & Botter, G. (2019). Flow dynamics at the continental scale: Streamflow correlation and hydrological similarity. *Hydrological Processes*, 33(4), 627–646. <https://doi.org/10.1002/hyp.13350>
- Blöschl, G., Sivapalan, M., Savenije, H., & Wagener, T. (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales*.
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A., Parajka, J., Merz, B., ... & Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, 573(7772), 108–111.
- Bracken, C., Rajagopalan, B., & Woodhouse, C. (2016). A Bayesian hierarchical nonhomogeneous hidden Markov model for multisite streamflow reconstructions. *Water Resources Research*, 52(10), 7837–7850. <https://doi.org/10.1002/2016WR018887>
- Bracken, C., Rajagopalan, B., Alexander, M., & Gangopadhyay, S. (2015). Spatial variability of seasonal extreme precipitation in the western United States. *Journal of Geophysical Research: Atmospheres*, 120(10), 4522–4533.
- Bracken, C., Rajagopalan, B., & Zagona, E. (2014). A hidden Markov model combined with climate indices for multidecadal streamflow simulation. *Water Resources Research*, 50(10), 7836–7846. <https://doi.org/10.1002/2014WR015567>

- Bourges, M., Mari, J.-L., & Jeannée, N. (2012). A practical review of geostatistical processing applied to geophysical data: methods and applications. *Geophysical Prospecting*, 60(3), 400–412. <https://doi.org/10.1111/j.1365-2478.2011.00992.x>
- Buckby, J., Wang, T., Zhuang, J., & Obara, K. (2020). Model Checking for Hidden Markov Models. *Journal of Computational and Graphical Statistics*, 29(4), 859-874.
- Burgdorf, A. M., Brönnimann, S., & Franke, J. (2019). Two types of North American droughts related to different atmospheric circulation patterns. *Climate of the Past*, 15(6), 2053-2065.
- Carrier, C., Kalra, A., & Ahmad, S. (2013). Using Paleo Reconstructions to Improve Streamflow Forecast Lead Time in the Western United States. *Journal of the American Water Resources Association*, 49(6), 1351–1366. <https://doi.org/10.1111/jawr.12088>
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, 1(2), 245-276.
- Cayan, D. R. (1996). Interannual climate variability and snowpack in the western United States. *Journal of climate*, 9(5), 928-948.
- Celeux, G., & Durand, J. B. (2008). Selecting hidden Markov model state number with cross-validated likelihood. *Computational Statistics*, 23, 541-564.
- Chen, F., Shang, H., Panyushkina, I. P., Meko, D. M., Yu, S., Yuan, Y., & Chen, F. (2019). Tree-ring reconstruction of Lhasa River streamflow reveals 472 years of hydrologic change on southern Tibetan Plateau. *Journal of Hydrology*, 572, 169–178. <https://doi.org/10.1016/J.JHYDROL.2019.02.054>
- Chen, J., Brissette, F. P., & Leconte, R. (2011). Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *Journal of Hydrology*, 401(3–4), 190–202. <https://doi.org/10.1016/J.JHYDROL.2011.02.020>
- Chang, K. T. (2009). Computation for bilinear interpolation. *Introduction to Geographic Information Systems*, 5th ed. New York, NY: McGraw-Hall. Print, 12.
- Cook, E. R., Meko, D. M., Stahle, D. W., & Cleaveland, M. K. (1999). Drought reconstructions for the continental United States. *Journal of Climate*, 12(4), 1145-1162.
- Cook, E. R., Seager, R., Cane, M. A., & Stahle, D. W. (2007). North American drought: Reconstructions, causes, and consequences. *Earth-Science Reviews*, 81(1-2), 93-134.
- Cook, E.R., Seager, R., Heim, R.R., Vose, R.S., Herweijer, C., and Woodhouse, C. 2010. Megadroughts in North America: Placing IPCC projections of hydroclimatic change in a long-term paleoclimate context. *Journal of Quaternary Science*, 25(1), 48-61. doi: 10.1002/jqs.1303
- Diffenbaugh, N. S., Swain, D. L., & Touma, D. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*, 112(13), 3931-3936.

- Doulatyari, B., Betterle, A., Radny, D., Celegon, E. A., Fanton, P., Schirmer, M., & Botter, G. (2017). Patterns of streamflow regimes along the river network: The case of the Thur river. *Environmental Modelling & Software*, 93, 42–58. <https://doi.org/10.1016/j.envsoft.2017.03.002>
- Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in bioinformatics*, 21(2), 553-565.
- Frederick, K.D. & P.H. Gleick. 1999. *Water and Global Climate Change: Potential Impacts on U.S. Water Resources*. Pew Center on Global Climate Change. Arlington , VA .
- Fortin, V., Perreault, L., & Salas, J. D. (2004). Retrospective analysis and forecasting of streamflows using a shifting level model. *Journal of hydrology*, 296(1-4), 135-163.
- Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Review Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol*, 27, 1547–1578. <https://doi.org/10.1002/joc.1556>
- Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2021). *An introduction to statistical learning: with applications in R*. Second edition. Springer.
- Groves, D. G., Yates, D., & Tebaldi, C. (2008). Developing and applying uncertain global climate change projections for regional water management planning. *Water Resources Research*, 44(12).
- Guo, Z., Tang, J., Tang, J., Wang, S., Yang, Y., Luo, W., & Fang, J. (2022). Object-based evaluation of precipitation systems in convection-permitting regional climate simulation over Eastern China. *Journal of Geophysical Research: Atmospheres*, 127(1), e2021JD035645.
- Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal drought prediction: advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1), 108-141.
- Hayes, M., Svoboda, M., Le Comte, D., Redmond, K. T., & Pasteris, P. (2005). Drought monitoring: New tools for the 21st century. *Drought and water crises: Science, technology, and management issues*, 53, 69.
- Heddinghaus, T. R., & Sabol, P. (1991). A review of the Palmer Drought Severity Index and where do we go from here. In *Proc. 7th Conf. on Applied Climatology* (pp. 242-246). Boston, MA: American Meteorological Society.
- Huang, Y., Liu, B., Zhao, H., & Yang, X. (2022). Spatial and Temporal Variation of Droughts in the Mongolian Plateau during 1959–2018 Based on the Gridded Self-Calibrating Palmer Drought Severity Index. *Water*, 14(2), 230.
- Hughes, J. P., & Guttorp, P. (1994). A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water resources research*, 30(5), 1535-1546. <https://doi.org/10.1029/93WR02983>

- Ho, M., Lall, U., & Cook, E. R. (2018). How wet and dry spells evolve across the conterminous United States based on 555 years of paleoclimate data. *Journal of Climate*, 31(16), 6633-6647.
- Ho, M., Lall, U., & Cook, E. R. (2016). Can a paleodrought record be used to reconstruct streamflow?: A case study for the Missouri River Basin. *Water Resources Research*, 52(7), 5195–5212. <https://doi.org/10.1002/2015WR018444>.Received
- Ho, M., Lall, U., Sun, X., & Cook, E. R. (2017). Multiscale temporal variability and regional patterns in 555 years of conterminous U.S. streamflow. *Water Resources Research*, 53(4), 3047–3066. <https://doi.org/10.1002/2016WR019632>
- Holsclaw, T., Greene, A. M., Robertson, A. W., & Smyth, P. (2017). Bayesian nonhomogeneous Markov models via Pólya-Gamma data augmentation with applications to rainfall modeling. *The Annals of Applied Statistics*, 11(1), 393-426.
- Keller, A. A., Garner, K. L., Rao, N., Knipping, E., & Thomas, J. (2022). Downscaling approaches of climate change projections for watershed modeling: Review of theoretical and practical considerations. *PLOS Water*, 1(9), e0000046.
- Kisilevich, S., Mansmann, F., Nanni, M., Rinzivillo, S. (2009). Spatio-temporal clustering. In: Maimon, O., Rokach, L. (eds) *Data Mining and Knowledge Discovery Handbook*. Springer, Boston, MA. [https://doi.org/10.1007/978-0-387-09823-4\\_44](https://doi.org/10.1007/978-0-387-09823-4_44)
- Kiewiet, L., Trujillo, E., Hedrick, A., Havens, S., Hale, K., Seyfried, M., ... & Godsey, S. E. (2022). Effects of spatial and temporal variability in surface water inputs on streamflow generation and cessation in the rain–snow transition zone. *Hydrology and Earth System Sciences*, 26(10), 2779-2796.
- Lehner, F., Wahl, E. R., Wood, A. W., Blatchford, D. B., & Llewellyn, D. (2017). Assessing recent declines in Upper Rio Grande runoff efficiency from a paleoclimate perspective. *Geophysical Research Letters*, 44(9), 4124-4133.
- Lins, H. F. (1997). Regional streamflow regimes and hydroclimatology of the United States. *Water Resources Research*, 33(7), 1655-1667.
- Malmgren, B. A., & Winter, A. (1999). Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network. *Journal of climate*, 12(4), 977-985.
- Marston, M. L., & Ellis, A. W. (2021). Delineating precipitation regions of the contiguous United States from cluster analyzed gridded data. *Annals of the American Association of Geographers*, 111(6), 1721-1739.
- Maurer, E. P., & Hidalgo, H. G. (2008). Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, 12(2), 551–563. <https://doi.org/10.5194/hess-12-551-2008>

- Mehrotra, R., & Sharma, A. (2005). A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences. *Journal of Geophysical Research: Atmospheres*, 110(D16).
- Merz, R., & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, 287(1–4), 95–123. <https://doi.org/10.1016/J.JHYDROL.2003.09.028>
- National Research Council. (2007). *Colorado River Basin water management: Evaluating and adjusting to hydroclimatic variability*. National Academies Press.
- Ogunrinde, A. T., Oguntunde, P. G., Olasehinde, D. A., Fasinmirin, J. T., & Akinwumiju, A. S. (2020). Drought spatiotemporal characterization using self-calibrating Palmer Drought Severity Index in the northern region of Nigeria. *Results in Engineering*, 5, 100088.
- Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences*, 11(5), 1633-1644.
- Pongracz, R., Bogardi, I., & Duckstein, L. (1999). Application of fuzzy rule-based modeling technique to regional drought. *Journal of Hydrology*, 224(3-4), 100-114.
- Rao, M. P., Cook, E. R., Cook, B. I., Palmer, J. G., Uriarte, M., Devineni, N., Lall, U., D'Arrigo, R. D., Woodhouse, C. A., Ahmed, M., Zafar, M. U., Khan, N., Khan, A., & Wahab, M. (2018). Six Centuries of Upper Indus Basin Streamflow Variability and Its Climatic Drivers. *Water Resources Research*, 54(8), 5687–5701. <https://doi.org/10.1029/2018WR023080>
- Razavi, T., & Coulibaly, P. (2013). *Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods*. [https://doi.org/10.1061/\(ASCE\)](https://doi.org/10.1061/(ASCE))
- Raziei, T., Bordi, I., & Pereira, L. S. (2008). A precipitation-based regionalization for Western Iran and regional drought variability. *Hydrology and Earth System Sciences*, 12(6), 1309-1321.
- Redmond, K. T., & Koch, R. W. (1991). Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices. *Water resources research*, 27(9), 2381-2399.
- Rice, J. L., Woodhouse, C. A., & Lukas, J. J. (2009). Science and Decision Making: Water Management and Tree-Ring Data in the Western United States. *JAWRA Journal of the American Water Resources Association*, 45(5), 1248–1259. <https://doi.org/10.1111/j.1752-1688.2009.00358.x>
- Robertson, A. W., Kirshner, S., & Smyth, P. (2003). *Hidden Markov models for modeling daily rainfall occurrence over Brazil*. Information and Computer Science, University of California.
- Salehnia, N., Hosseini, F., Farid, A., Kolsoumi, S., Zarrin, A., & Hashemina, M. (2019). Comparing the Performance of Dynamical and Statistical Downscaling on Historical Run Precipitation Data over a Semi-Arid Region. *Asia-Pacific Journal of Atmospheric Sciences*, 1–13. <https://doi.org/10.1007/s13143-019-00112-1>

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American statistical association*, 63(324), 1343-1372.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., Mcdonnell, J. J., Mendiondo, E. M., O'connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., & Zehe, E. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, 48(6), 857–880. <https://doi.org/10.1623/hysj.48.6.857.51421>

Son, R., Wang, S. S., Kim, S. H., Kim, H., Jeong, J. H., & Yoon, J. H. (2021). Recurrent pattern of extreme fire weather in California. *Environmental Research Letters*, 16(9), 094031.

Sunyer, M. A., Madsen, H., & Ang, P. H. (2012). A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change. *Atmospheric Research*, 103, 119–128. <https://doi.org/10.1016/J.ATMOSRES.2011.06.011>

Swain, S., & Hayhoe, K. (2015). CMIP5 projected changes in spring and summer drought and wet conditions over North America. *Climate Dynamics*, 44(9), 2737-2750.

Timilsena, J., & Piechota, T. (2008). Regionalization and reconstruction of snow water equivalent in the upper Colorado River basin. *Journal of Hydrology*, 352(1-2), 94-106.

Tisseuil, C., Vrac, M., Lek, S., & Wade, A. J. (2010). Statistical downscaling of river flows. *Journal of Hydrology*, 385(1–4), 279–291. <https://doi.org/10.1016/J.JHYDROL.2010.02.030>

USGS Publication “What is a Desert.” Accessed September 21, 2022, <https://pubs.usgs.gov/gip/deserts/what/#:~:text=Deserts%20are%20natural%20laboratories%20in,that%20were%20exposed%20by%20erosion.>

Vallam, P., & Qin, X. S. (2017). Climate change impact assessment on flow regime by incorporating spatial correlation and scenario uncertainty. *Theoretical and Applied Climatology*, 129(1–2), 607–622. <https://doi.org/10.1007/s00704-016-1802-1>

Vinnarasi, R., & Dhanya, C. T. (2022). Time-varying Intensity-Duration-Frequency relationship through climate-informed covariates. *Journal of Hydrology*, 604, 127178.

Voltaire, A., E. Sanchez-Gomez, D. Salas y Méliá, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, M. Chevallier, M. Déqué, J. Deshayes, H. Douville, E. Fernandez, G. Madec, E. Maisonnave, M.-P. Moine, S. Planton, D. Saint-Martin, S. Szopa, S. Tyteca, R. Alkama, S. Belamari, A. Braun, L. Coquart, F. Chauvin (2011) : The CNRM-CM5.1 global climate model : description and basic evaluation, *Clim. Dyn.*, accepted, DOI:10.1007/s00382-011-1259-y.

Vrac, M., & Naveau, P. (2007). Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resources Research*, 43(7). <https://doi.org/10.1029/2006WR005308>

Wahl, E.R. and J.E. Smerdon. 2012. Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noise-only predictors. *Geophys. Res. Lett.*, 39, L06703, doi:10.1029/2012GL051086

Wahl, E. R., Zorita, E., Diaz, H. F., & Hoell, A. (2022). Southwestern United States drought of the 21st century presages drier conditions into the future. *Communications Earth & Environment*, 3(1), 202.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

White, D., Richman, M., & Yarnal, B. (1991). Climate regionalization and rotation of principal components. *International Journal of Climatology*, 11(1), 1-25.

Woldemeskel, F. M., Sharma, A., Sivakumar, B., & Mehrotra, R. (2016). Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. *Journal of Geophysical Research: Atmospheres*, 121(1), 3–17. <https://doi.org/10.1002/2015JD023719>

Woodhouse, C. A., & Lukas, J. J. (2006a). Drought, Tree Rings and Water Resource Management in Colorado. *Canadian Water Resources Journal*, 31(4), 297–310. <https://doi.org/10.4296/cwrj3104297>

Woodhouse, C. A., & Lukas, J. J. (2006b). Multi-Century Tree-Ring Reconstructions of Colorado Streamflow for Water Resource Planning. *Climatic Change*, 78(2–4), 293–315. <https://doi.org/10.1007/s10584-006-9055-0>

Xu, C.-Y. (1999). *From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches*.

Xu, Z., Han, Y., & Yang, Z. (2019). Dynamical downscaling of regional climate: A review of methods and limitations. *Science China Earth Sciences*, 62(2), 365–375. <https://doi.org/10.1007/s11430-018-9261-5>

Yu, Y., Shao, Q., & Lin, Z. (2018). Regionalization study of maximum daily temperature based on grid data by an objective hybrid clustering approach. *Journal of Hydrology*, 564, 149-163.

Zucchini, W., & MacDonald, I. L. (2009). *Hidden Markov models for time series: an introduction using R*. Chapman and Hall/CRC.

## Appendix A

### *A.1 K-means Clustering*

K-means cluster analysis is also used to reduce dimensionality but not pursued in this study due to having a major drawback in terms of failing to capture the PMDI variability. The estimated cluster centers are found to be close to overall mean of the variables (which is close to zero). The methodology and results for the k-means cluster analysis are given in this section.

#### *A.1.1 Methodology*

Cluster analysis allows the analyst to separate data into groups with similar properties and can identify key features in large data sets (Wilks, 2011). It allows the analyst to focus on a higher-level representation of the data. In addition, it analyzes and explores a dataset to associate objects in groups that have common characteristics (Kisilevich, 2009). Two main clustering algorithms, hierarchical and nonhierarchical, are used for gridded data. In this study, a nonhierarchical clustering approach, k-means clustering is used to combine gridded data based on their temporal similarity. K-means clustering is a centroid-based cluster method which it starts by computing the centroids for each cluster and then calculates the distances between current data vector and each of centroids (Wilks, 2011). The k-means method is a very good approach identifying patterns related to mean behavior of Gaussian or Gaussian mixture data (Wilks, 2011; Bracken et. al., 2015). K-means clustering has been applied successfully to hydroclimatic variables including drought indices (Huang et. al., 2021), precipitation (Marston and Ellis, 2021), and temperature records (Yu and Lin, 2018).

The optimal number of clusters is determined with various evaluation metrics, including an elbow method with difference in the within cluster sum of square errors, a silhouette index with the average distance between clusters, gap statistics with difference in difference, a Calinski-Harabasz index with the degree of dispersion between clusters, and qualitative analysis on post-visualization of clusters. The optimal number of clusters are identified for PMDI and temperature reconstructions based on majority of the metrics if there is disagreement across metrics.

#### *A.1.2. Results*

The optimal number of groups formed by the k-means algorithm for the study area was determined based on a series of methods described above (Silhouette, Scree plot, Calinski criterion, etc.). Silhouette and Scree plots for the PMDI reconstructions in the western U.S. are given in Figure A.1 as an example.

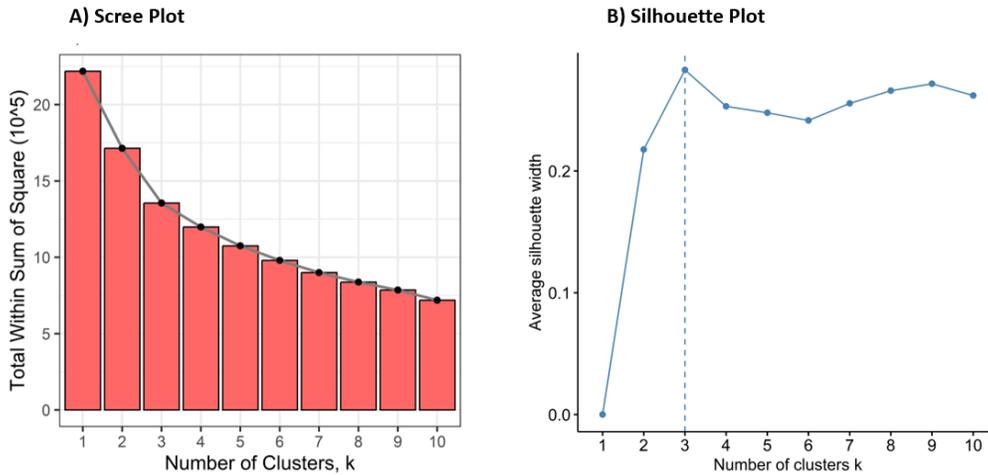


Figure A.1. K-Means clustering selection metrics for the PMDI reconstructions: A) Scree plot and B) Silhouette plot.

Three clusters are identified for PMDI reconstructions, and two clusters are determined for temperature reconstructions. The spatial distribution of the center of the grid cells of clusters for PMDI and temperature reconstructions are shown in Figure A.2.

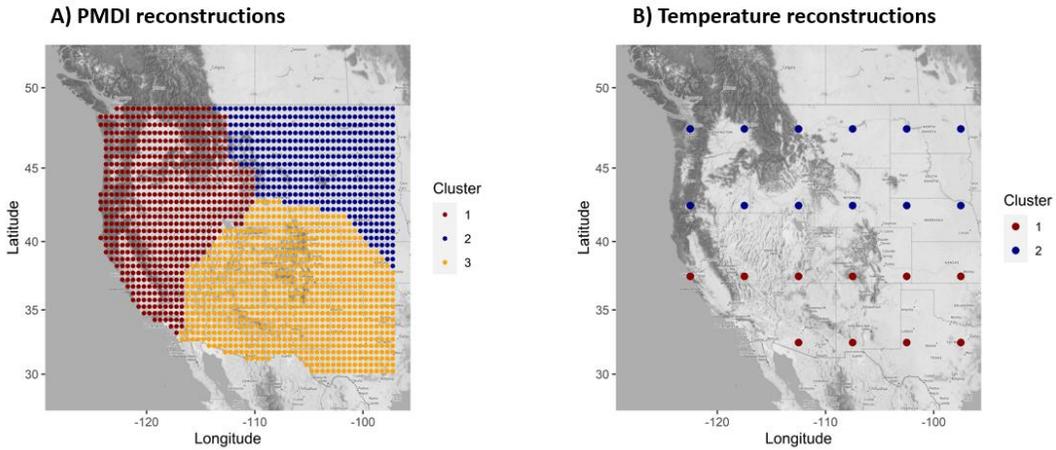


Figure A.2. Spatial Distribution of A) PMDI and B) temperature reconstruction clusters in the western U.S.

The NHMM model with k-means clustering provides insight into dry and wet events via independently Gaussian distributed variables with mean and standard deviation where both mean and standard deviation are unique to each cluster for a given state. For example, the driest state is identified as state-5 with the lowest mean of 0.033, the wettest state is determined as state-6 with the highest mean of 0.146 for the first cluster of the western U.S. for year 1500. (Figure A.3).

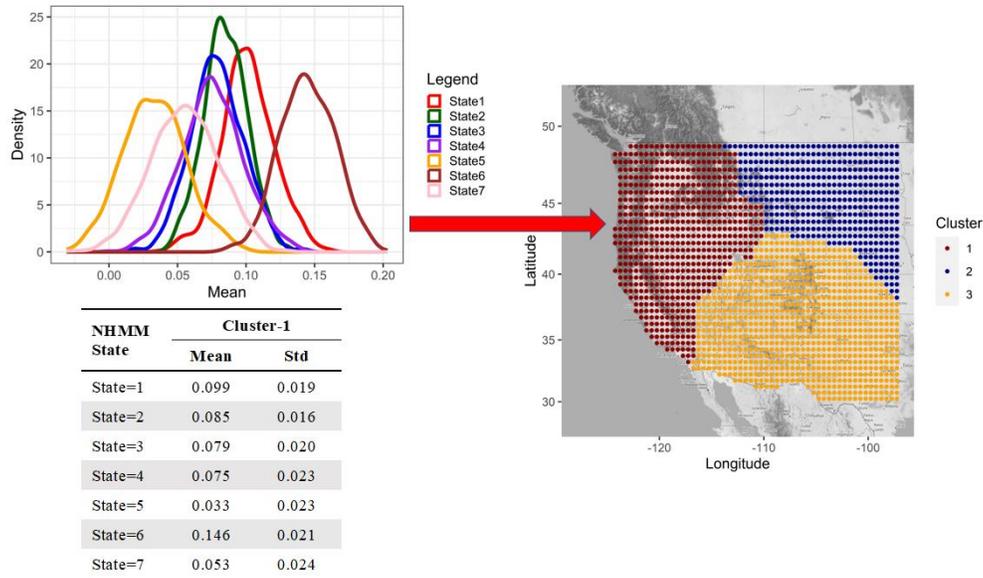


Figure A.3. State Distributions for the first cluster of the western U.S. with year of 1500.

Figure A.4 shows PMDI and temperature reconstruction anomaly time series together with most likely hidden states for the first cluster during training period. As seen from section A of the figure, the estimated cluster centers of PMDI are found to be close to overall mean of the variables (which is close to zero). This result is the same for all clusters. Positive PMDI represents wet conditions while negative PMDI represents dry conditions. States represent wet, average, and dry conditions.

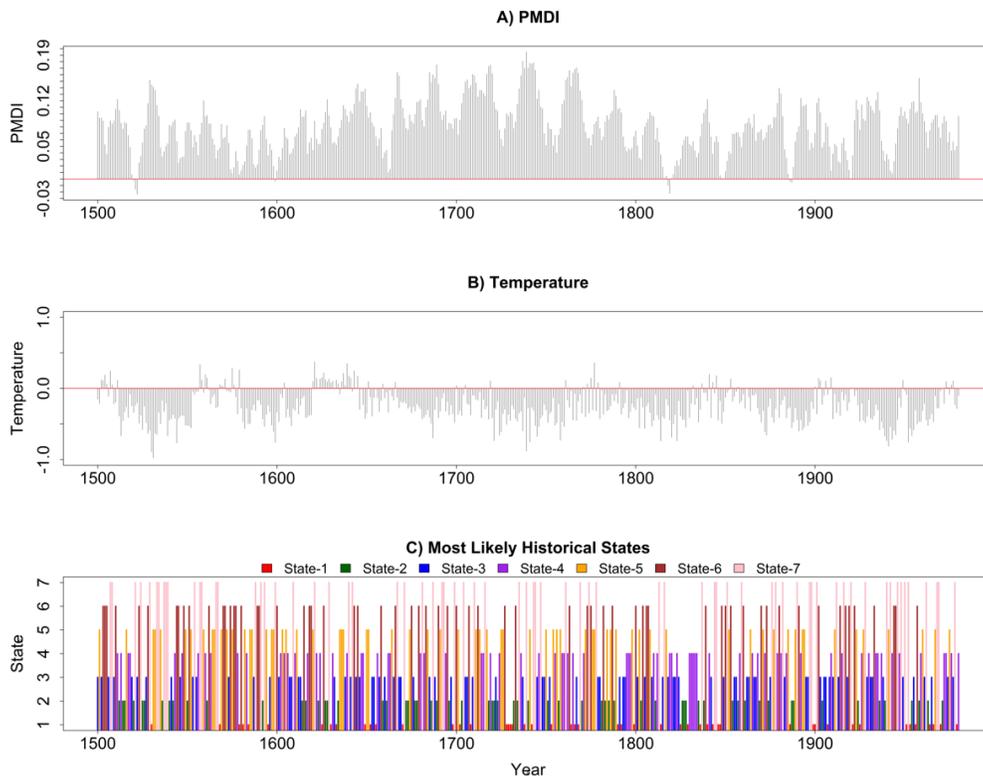


Figure A.4. A) Annual mean PMDI reconstruction time series for the first cluster, B) Annual mean temperature reconstruction anomaly time series for the first cluster, and C) Most likely states over the western U.S.

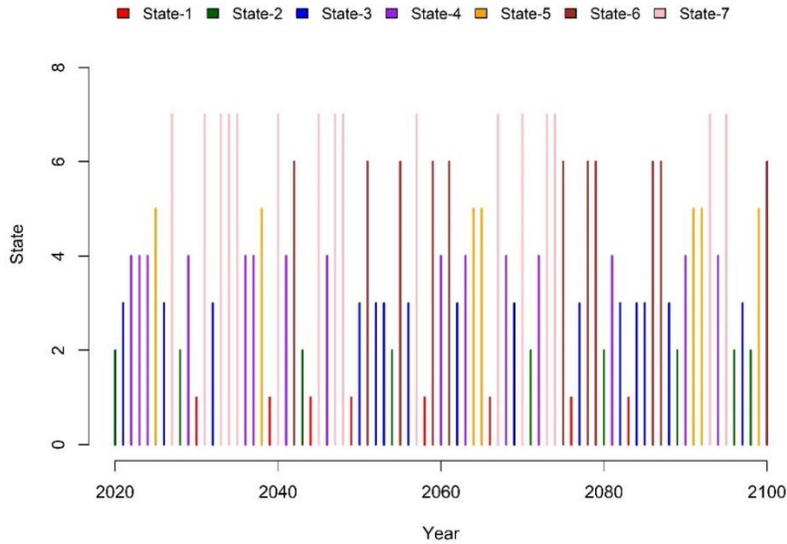


Figure A.5 One stochastic scenario of system states in the western U.S. (Scenario – 1).

K-means clustering analysis results in having a major drawback in terms of failing to capture the PMDI variability over the western U.S. The study region exhibits greater spatio temporal rainfall variability, which is closely linked to PMDI. The large spatial extent of the clusters may lead to failure capturing the variability due to having extreme values with no similar high- or low-value neighbors, as stated in Wong (2021).