**Climate Dynamics manuscript No.**
(will be inserted by the editor)

# Pattern-based downscaling of snowpack variability in the western United States

**Nicolas Gauthier · Kevin J. Anchukaitis · Bethany Coulthard**

Received: date / Accepted: date

**Abstract** The decline in snowpack across the western United States is one of the most pressing threats posed by climate change to regional economies and livelihoods. Earth system models are important tools for exploring past and future snowpack variability, yet their coarse spatial resolutions distort local topography and bias spatial patterns of accumulation and ablation. Here, we explore pattern-based statistical downscaling for spatially-continuous interannual snowpack estimates. We find that a few leading patterns capture the majority of snowpack variability across the western US in observations, reanalyses, and free-running simulations. Pattern-based downscaling methods yield accurate, high resolution maps that correct mean and variance biases in domain-wide simulated snowpack. Methods that use large-scale patterns as both predictors and predictands perform better than those that do not and all are superior to an interpolation-based "delta change" approach. These findings suggest that pattern-based methods are appropriate for downscaling interannual snowpack variability and that using physically meaningful large-scale patterns is more important than the details of any particular downscaling method.

**Keywords** snow water equivalent · empirical orthogonal functions · canonical correlation analysis · teleconnections · water resources

This research was supported by the National Science Foundation Paleo Perspectives on Climate Change (P2C2) grant AGS-1803995.

N. Gauthier
School of Geography, Development & Environment and Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ E-mail: ngauthier@arizona.edu

K. J. Anchukaitis
School of Geography, Development & Environment and Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ

B. Coulthard
Department of Geoscience, University of Nevada Las Vegas, Las Vegas, NV

## 1 Introduction

The decline in snowpack across the western United States is one of the most pressing threats posed by climate change to regional economies and livelihoods (Mankin and Diffenbaugh, 2015; Mote et al, 2018; Xiao et al, 2018; Huning and AghaKouchak, 2020). Spring snowmelt is critical for regional water managers—more than half of annual runoff in the western US derives from snowpack (Li et al, 2017). Snow plays a central role in local and regional climates and ecosystems, from its cooling effect on temperatures to its modulation of the timing and intensify of streamflow and soil moisture anomalies (Walsh et al, 1982; Marks and Dozier, 1992; Bales et al, 2006; Maurer and Bowling, 2014; Li et al, 2017). The observed decline in snowpack is the result of several interacting factors including shifts in the timing and intensity of seasonal precipitation and temperature patterns, each of which are exacerbated by warming temperature trends and the attendant changes in accumulation and ablation (Pierce et al, 2008; Kapnick and Hall, 2012; Pederson et al, 2013; Klos et al, 2014; Xiao et al, 2018). These snowpack deficits are of a magnitude and extent unprecedented in the observational period (McCabe and Wolock, 2009; Mote et al, 2018; Schoenemann et al, 2020) and are expected to worsen in the future (Fyfe et al, 2017; Marshall et al, 2019; Siler et al, 2019).

Yet it remains difficult to observe snowpack uniformly across large spatial domains. Spatially-continuous high-resolution maps of snowpack are therefore a challenge to produce, particularly in areas with complex terrain (Erickson et al, 2005; Meromy et al, 2013). Different sensor types and measurement strategies focus on distinct—if related—facets of the system, such as snow water equivalent (SWE), snow-covered area, and snow depth. Each has unique uncertainties, coverage, and observational spans, making them a challenge to integrate (Dozier et al, 2016; Dong, 2018). In most locations the observational record only extends for a few decades into the past (e.g. Serreze et al, 1999), making it difficult to place observed variability in a long-term context.

An array of modeling approaches provides ways to estimate gaps in the observational record and produce continuous spatiotemporal data. From standalone hydrological bucket models to the complex land-surface components of Earth system models, snowpack simulations attempt to capture the interacting drivers of snowpack variability across spatial and temporal scales. These models allow for assessments of the mechanistic uncertainty of these drivers and uncertainty in their observations (Clark et al, 2011). Even simple models provide useful information for constraining noisy observations (Broxton et al, 2016a). Although the skill of current-generation snow models is high overall, issues remain in the representation of processes like ablation at near-freezing temperatures (Rutter et al, 2009; Broxton et al, 2016b; Krinner et al, 2018). Regional and global snow models must run on daily to sub-daily time scales, so a reduction in spatial resolution may be required to minimize computational costs. This tradeoff makes accurate spatial modeling of snowpack difficult, even when the underlying process models are physically appropriate.

Snow accumulation and ablation is sensitive to local topography, particularly in the mountainous regions that receive the most snowfall (Anderson et al, 2014; Tennant et al, 2017; Jennings and Molotch, 2019). The resolution of most simulations smooth this topography, eliminating mountain peaks and introducing temperature biases that prevent snow from accumulating where it otherwise would (Rhoades et al, 2018). The tendency for snow models to underpredict accumulated SWE has been well documented. Xu et al (2019) showed that increasing model resolution from 0.44°to 0.11°increases the accuracy of simulated SWE by 35%. Such low-snow biases in regional and global snow simulations preclude their use by local water managers without corrections to this fundamental scale mismatch. Some form of downscaling is required to estimate fine-resolution snowpack maps from coarser-resolution simulation outputs (McGinnis, 1997; Pons et al, 2010; Tryhorn and Degaetano, 2013). However, this is increasingly accomplished via an additional high-resolution regional climate model or by forcing a hydrological model with atmospheric forcing data downscaled by constructed analogue methods, both of which require data on hourly to daily time scales, making them computationally infeasible for assessing variability on time horizons greater than a few decades (Rhoades et al, 2018; Chegwidden et al, 2019; Fiddes et al, 2019; Ikeda et al, 2021).

Non-local "pattern-based" statistical downscaling methods are an effective alternative to quickly generate fine-scale, long-term ensembles from existing coarse-resolution climate model simulations. Pattern-based methods decompose observed and simulated climate fields into a limited number of spatiotemporal patterns or "modes of variability," finding statistical relationships that translate one set of modes into the other (Bretherton et al, 1992; Tippett et al, 2008; Simon et al, 2013; Maraun and Widmann, 2018). Because they find associations between internally-consistent predictor and predictand fields, pattern-based statistical methods share some benefits with more computationally expensive dynamic downscaling methods that preserve the physical consistency of the simulated climate fields. These methods are "non-local" in that they focus on associations between large-scale patterns, rather than local associations between an observed location and the overlapping simulation grid cell. The simulation grid cell that best captures the observed variability at a given location is often *not* the corresponding local grid cell (van den Dool et al, 2000; Maraun and Widmann, 2015; Nicholson et al, 2019). While local mean conditions reflect local terrain, year-to-year departures from the mean often reflect teleconnections to remote, large-scale atmosphere-ocean variability (van den Dool et al, 2000; Hewitt et al, 2018). Anchoring the downscaling process in these large-scale physical mechanisms leads to a higher signal to noise ratio (Benestad et al, 2015), ensuring the estimated statistical relationships are internally consistent and likely to remain stable over time.

Here, we explore pattern-based statistical methods for downscaling interannual variability in March mean SWE across the western United States. We find that a few leading modes—present in observations, simulations, and reanalyses—capture the majority of snowpack variability in this domain. We compare several related regression methods for finding associations between observed and simulated patterns and show that even simple linear models perform well under cross validation. These methods yield accurate high resolution maps that correct mean and variance biases in domain-wide simulated SWE. Methods that use large-scale patterns as both predictors and predictands perform better than those that use those patterns on only one side of the regression equation, and all pattern-based methods are superior to a local "delta change" approach. These findings suggest that pattern-based methods are indeed appropriate for downscaling interannual snowpack variability, and that employing physically-meaningful large-scale patterns is more important for accuracy than the details of any particular downscaling method. Our findings here demonstrate the utility of applying these approaches where more computational- or data-intensive methods are impractical, including paleoclimate modeling and data assimilation.

## 2 Data

### 2.1 Observations

We focus on a domain between 125°W–102°W and 31°N–49°N, covering the western US states of Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. Observed March SWE was calculated from the University of Arizona (UA) Daily 4km SWE data product, a gridded record of daily SWE and snow depth for water years 1982-2017 at 4km resolution across the conterminous US (Broxton et al, 2019). March mean SWE has been shown to approximate the more commonly used April 1st SWE measure, but is less sensitive to sampling variability than a single daily value (Mankin and Diffenbaugh, 2015; Ye, 2019). The UA SWE data were based on a simple ablation and accumulation model driven by gridded daily PRISM temperature and precipitation fields (Daly et al, 2008), rescaled by relative anomalies from thousands of *in situ* observations from the SNOTEL and COOP networks (Broxton et al, 2016a; Zeng et al, 2018). We also acquired the raw PRISM temperature and precipitation fields to assess local relationships between SWE accumulation and seasonal hydroclimate variability.

### 2.2 Reanalyses and Simulations

Modeled SWE for the downscaling experiments was derived from the CERA 20th century (CERA-20C) reanalysis product (variable name *SD*) (Laloyaux et al, 2018). CERA-20C is a long-term reanalysis product that uses the European Centre for Medium-Range Weather Forecast (ECMWF) system spanning 1901-2010 at six-hourly temporal resolution and ~1°spatial resolution. It assimilates sea level pressure pressure and ocean temperature observations from across this period in order to avoid temporal inconsistencies from the later introduction of, for example, satellite observations. We also acquired monthly sea surface temperatures and 500mb geopotential heights from the same reanalysis to assess large-scale atmosphere-ocean teleconnections. We used the means of the 10-member ensemble for all analyses as the individual ensemble members showed few major differences during the period of interest.

As a preliminary evaluation of whether these methods could be applied to free-running paleoclimate model simulations, we also analyzed outputs from the CCSM4 Last Millennium simulation (Landrum et al, 2013) and the CESM Last Millennium Ensemble (Otto-Bliesner et al, 2016), their associated 20th century extensions (variable name *H2OSNO*, CMIP5 standard name *SNW*), and version 3 of the NOAA-CIRES-DOE 20th Century Reanalysis (20CRv3) (variable name *WEASD*) (Slivinski et al, 2020) in order to assess modes of snowpack variability in free-running Earth system models of different native resolutions ($\sim$1°and $\sim$2°) and reanalysis data from different modeling groups, respectively. Herein, we collectively refer to both reanalyses and free-running climate models as "simulations" for simplicity.

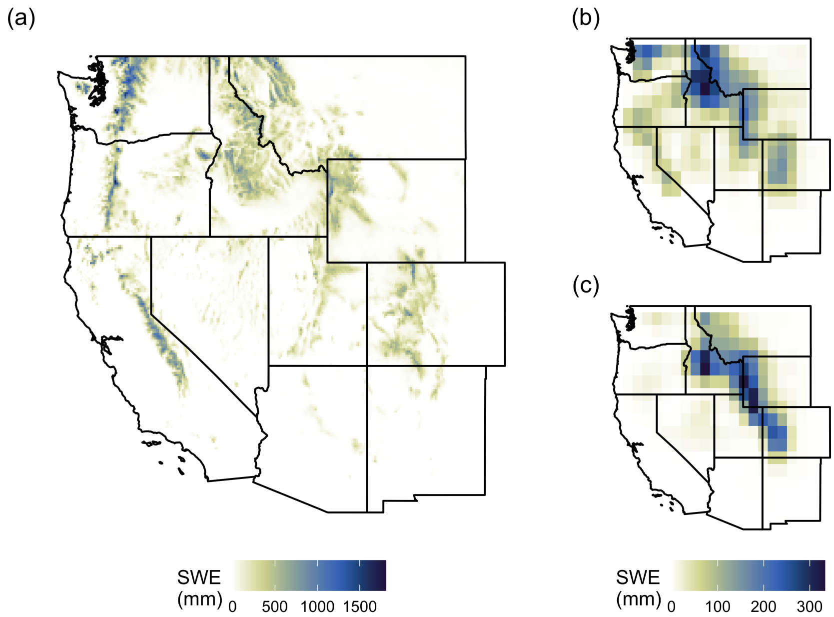


**Fig. 1** Mean March snow water equivalent (SWE) in mm from A) UA 4km daily SWE observations (Broxton et al, 2019), B) CERA-20C reanalysis (Laloyaux et al, 2018), C) CCSM4 Last Millennium simulation (Landrum et al, 2013). Note the scale of the observations differs from the simulations by nearly an order of magnitude due to differences in model resolution.

### 2.3 Preprocessing

Both observed and simulated data were truncated to the overlapping period of 1982-2010 and aggregated from daily to monthly timescales by calculating the average March SWE value for each grid cell and year (Figure 1). We used bilinear interpolation to resample each of the large-scale simulation outputs to a common 1°grid. We also resampled the 4km snow observations to an 8km grid to decrease computational costs without degrading the high-resolution spatial signal. Grid cells that experienced no SWE accumulation throughout the observational period were masked from successive analyses.

## 3 Methods

### 3.1 Estimating modes of snowpack variability

We isolated key modes of observed snowpack variability using principal components analysis (PCA). The observed and simulated data were area weighted to prevent undue influence from grid cells at higher latitudes by multiplying the observations at each grid cell by the square root of of the cosine of the cell's latitude in radians (Livezey and Smith, 1999). We calculated interannual SWE anomalies by mean-centering the data before analysis. We do not use standardized or detrended anomalies in order to preserve spatial patterns of variance across the field (Zeng et al, 2018).

The PCA results in a set of orthogonal principal component time series or "amplitudes," eigenvalues representing the variation accounted for by each amplitude time series, and eigenvectors or "empirical orthogonal functions" (EOFs) mapping the amplitude time series back onto the original spatial grid. We standardized the PC amplitudes to unit variance and reweighted the eigenvectors by the square root of their corresponding eigenvalues to give higher weight to the leading spatial modes (Hannachi et al, 2007). Thus, the original dataset could be reconstructed by multiplying each amplitude time series by its corresponding EOF spatial pattern, summing the results to get SWE anomalies, and adding in the sample mean of the grid cell. Using only a subset of these spatiotemporal patterns to reconstruct the original SWE field effectively removes "noise" associated with the higher order modes, limiting the data to a subspace representing only the most important axes of variation. The truncation level $k$ for each field was selected by cross validation (see section 3.3).

We used several techniques to examine the leading spatiotemporal modes. We visualized the EOF modes by calculating the Pearson correlation coefficient between each PC amplitude time series and each grid cell's original time series. We explored potential atmosphere-ocean teleconnections by calculating the correlation between each PC amplitude and average October-March global sea surface temperatures (SSTs) and 500mb geopotential heights from the CERA-20C reanalysis (Laloyaux et al, 2018) and regional temperature and precipitation observations from PRISM (Daly et al, 2008), assessing statistically significant correlations using the false discovery rate (Wilks, 2006, 2016). We also applied a varimax rotation to the leading PCs to examine regional response patterns (Richman, 1986), although unrotated PCs were used for downscaling due to their favorable statistical properties and similarity to the rotated PCs.

Although we attempted to find physically-meaningful patterns where they were present, we did not consider the lack of physical interpretation to be a criterion for excluding a particular mode from the downscaling model. We ensured only that the retained modes *collectively* reflected large-scale atmosphere-ocean variability. In other words, the choice of truncation level $k$ and the combined set of coupled patterns were more important to our downscaling process than the physical interpretation of any particular mode.

### 3.2 Pattern-based downscaling

Pattern-based downscaling models use some combination of observed and simulated PC time series to predict one climate field from another. There are multiple statistical methods capable of doing so, many of which are variants on multiple linear regression (Bretherton et al, 1992; Tippett et al, 2008). They generally differ in whether they maximize explained variance in the observations as opposed to the shared variance between observations and simulations, and whether they use PCs as predictors, predictands, or both (Table 1). We compared four downscaling methods that spanned this methodological spectrum along with an additional "local" null model.

**Table 1** Pattern-based downscaling methods: canonical correlation analysis (CCA), principal components regression (PCR), principal components regression via generalized additive models (PCR-GAM), and empirical orthogonal teleconnections (EOT). Either the predictors (x), predictands (y), or both are subjected to PCA prefiltering prior to downscaling. Asymmetric models seek to explain variance of the predictands while symmetric models explain the shared correlation. Cross-validated performance metrics for the best-performing model of each class are the space-time root mean square error and the Pearson correlation between observed and simulated domain-wide SWE. The additive delta change approach using bilinearly interpolated anomalies is also including here as a local baseline for the nonlocal downscaling approaches.

| Method | PCA Prefiltering | Symmetric | RMSE | Correlation |
|---|---|---|---|---|
| CCA | $x$, $y$ | yes | 41.4 | 0.940 |
| PCR | $x$, $y$ | no | 43.1 | 0.949 |
| PCR-GAM | $x$, $y$ | no | 42.7 | 0.932 |
| EOT | $y$ | no | 48.5 | 0.918 |
| DELTA | *none* | no | 53.2 | 0.912 |

Canonical correlation analysis (CCA) is one of the most common approaches to coupled pattern analysis (Maraun and Widmann, 2018). It yields a set of patterns that maximizes the shared correlation between the predictor and predictand fields (Tippett et al, 2008). We applied CCA to the leading predictor and predictand modes of variability to regularize the model and make it computationally tractable (Barnett and Preisendorfer, 1987; Bretherton et al, 1992; Benestad, 2001; Tippett et al, 2008). Downscaling models are prone to overfitting on shorter calibration windows, so this PCA prefiltering step increases the signal-to-noise ratio to ensure the resulting patterns are statistically robust.

Principal components regression (PCR) is a similar method that uses the PC time series in independent multiple linear regressions. Traditional PCR fits a different model to the predictor PCs for each predictand grid cell, although here we take the more efficient approach of using predictand PCs directly (Benestad et al, 2015). Because the PC time series are mutually uncorrelated each predictand PC can be modeled independently and there is no concern of multicollinearity. PCR is asymmetric in that it only explains the variance of the predictands, contrary to CCA, although both methods are linear and are equivalent under certain conditions (Tippett et al, 2008). We also tested a nonlinear variant of PCR which replaces the linear models with penalized piecewise polynomials estimated in a generalized additive model (PCR-GAM).

Empirical orthogonal teleconnections (EOT) finds a set of grid cells that explain the most variance in the observation domain by fitting a linear model between all pairs of predictor and predictand grid cells (van den Dool et al, 2000; Appelhans et al, 2015). The simulation grid cell that predicts the most variance in all of the predictand grid cells is selected as the first pattern. Then the algorithm is run again on the residuals from the regressions on the first pattern, and the process is repeated until a set number of patterns is reached. EOT yields more localized spatial patterns, similar to rotated EOFs, than methods that use predictor and predictand PCs directly. Although PCA can be used to denoise both fields prior to the analysis, EOT focuses on the grid-cell level time series and is not constrained to fit the large scale patterns used by CCA and PCR.

We compared these non-local pattern-based techniques to a simple null model using interpolated simulation anomalies. This “delta change” model involved calculating the yearly simulated SWE anomalies relative to the mean fields, bilinearly interpolating the low-resolution simulated anomalies to the high resolution of the observations, and combining the interpolated anomalies with the high resolution observed mean. We tested delta change models with both additive and multiplicative anomalies by either subtracting the simulated mean and adding the observed mean or by dividing by the simulated mean and multiplying by the observed mean, respectively. While conceptually similar to the pattern-based methods, the delta change approach uses only local information and cannot correct any spatial biases caused by the smoothed topography of the simulation. We used these model to assess the added value of the non-local downscaling approaches relative to these common local methods.

All downscaling methods were implemented in R version 4.0.3 (R Core Team, 2020) using the packages `stars`, `tidyverse`, `mgcv`, `remote`, and `MuMIN`

(Wood, 2006; Appelhans et al, 2015; Wickham et al, 2019; Bartoń, 2020; Pebesma, 2021). Code for reproducing the main analysis and figures is available at https://github.com/nick-gauthier/tidyEOF.

## 3.3 Cross Validation

Each of the pattern-based downscaling methods required the number of coupled patterns to be defined by a hyperparameter $k$. The methods that used PCA prefiltering also required selection of a truncation level for the predictor and predictand PCs. We used a five-fold cross validation routine to tune the hyperparameters of each model, fitting and predicting from models with all possible combinations of up to ten predictor patterns $k_x$, predictand patterns $k_y$, and coupled patterns $k_{xy}$, with the constraint that $k_{xy} \leq min(k_x, k_y)$.

We divided the 29-year calibration period into five contiguous folds, four of which contained six years and one of which contained five. We held out one fold at a time, fitting each model and parameter combination on the remaining folds and using them to predict the held out fold. The entire modeling workflow—anomaly calculation, PCA truncation, and model fitting—was repeated for each training and testing fold independently to prevent leaking information among the folds (Van Den Dool, 1987; Livezey and Smith, 1999; Smerdon et al, 2010). We repeated this process until each fold had been used four times for training and once for testing, after which we combined the test folds into a single 29 year sequence from which we calculated the prediction error against the observed sequence. It is often preferable to use a nested cross validation routine when doing model selection and performance assessment simultaneously, but we did not do so in this case because our sample size was limited and the different models were of broadly the same type with a low number of similar hyperparameters (Wainer and Cawley, 2018).

We used two metrics to assess the skill of each model and parameter combination. First we examined the correlation between the observed and predicted domain-wide total SWE time series. We calculated total domain SWE by multiplying each SWE value by the area of its grid cell and summing the result. We then assessed the local spatial skill of the downscaled product by calculating the total space-time root mean square error (RMSE) between all observed and predicted grid cells. We selected the models and parameter combinations that maximized domain-wide correlation and minimized RMSE under cross validation, and refit the best performing model to the entire data series. We compared the predictions from this final model to the raw CERA-20C reanalysis to assess the added value of downscaling for correcting mean and variance biases in domain-wide SWE. We demonstrated the spatial skill of the model by comparing the spatial anomalies of observed, reanalysis, and downscaled fields during a known extreme year. To test its sensitivity to recent warming trends, we refit the best model holding out the years with the top 20% warmest October-March average temperatures in the PRISM observations. We also compared these CERA-20C based reconstructions to models using the NOAA-CIRES-DOE 20CRv3 reanalysis (Slivinski et al, 2020) as an alternative predictor to assess the sensitivity of the outputs to the specific reanalysis methodology.

A downscaling model trained on reanalysis data must also be able to make predictions from unseen, free-running simulations to make skillful climate-change impact assessments beyond the observational period (Maraun and Widmann, 2018). As a proof-of-concept of the generalizability of the final model and EOF patterns, we used it to downscale additional 300-year simulated snowpack sequences by projecting data from the CCSM4 and CESM Last Millennium simulations onto the reanalysis PC patterns. As these free-running simulations were not constrained to match the year-to-year evolution of the observations as were the reanalyses, the added value of downscaling was assessed through improvements in the mean and variance biases on a 50-year distributional basis.

## 4 Results

*A limited set of climate modes explain the majority of observed and simulated March SWE variance.* Four spatiotemporal patterns explain 76% of the observed variance in March snowpack over the western United States (Figure 2a). The leading ten patterns explain nearly 90% of the observed variance. These patterns represent recurring modes of spatiotemporal variability and are an efficient means of capturing the high dimensional spatiotemporal snowpack field in a limited subspace of patterns.

Similar patterns are found in coarse-resolution simulations. The leading 10 PCs of the 110 year CERA-20C reanalysis explain 96% of the variance in simulated snowpack, and the leading four explain 89% of the variance (Figure 2b). These reanalysis PCs are associated with the same broad spatial patterns as the observed PCs, but the longer sample windows allows for greater separation between the leading modes than with the 36 year observational record.

*Large-scale snow patterns reflect orography and atmosphere-ocean variability.* The spatial EOF patterns associated with the leading snowpack PCs exhibit clear relationships to regional precipitation and temperature (Figure 3) as well as global pressure systems and sea surface temperatures (Figure 4). EOF/PC1 is a domain-wide signal with high loadings in the Rocky Mountains, Sierra Nevada, and Cascade ranges. It is associated with simultaneous cold and wet conditions (or vice versa) over the domain and anomalous pressure systems over northwestern North America. EOF/PC2 exhibits a north-south dipole pattern with opposite-sign loadings in the Cascades and northern Rockies and the Sierra Nevada and southern Rockies, respectively. Unlike EOF1, this pattern is associated the precipitation, not temperature, anomalies over the domain and a far more zonal geopotential height anomaly over North America. EOF3 is localized to the Rocky Mountains and is associated with domain-wide temperature anomalies and geopotential and SST dipoles over the north

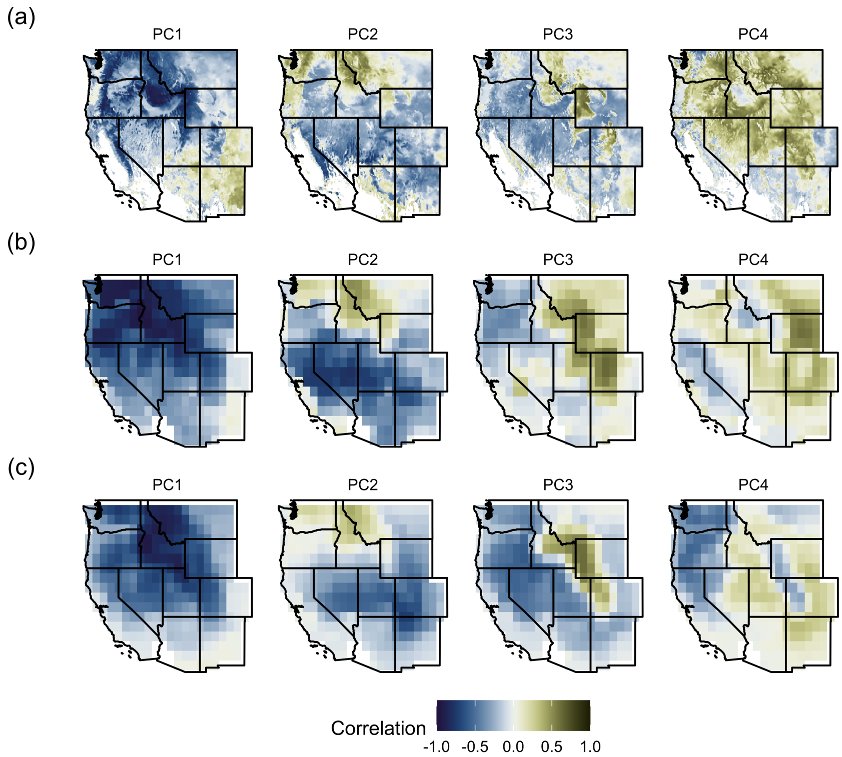


**Fig. 2** The leading four EOF spatial patterns expressed as the Pearson correlation coefficient between each PC time series and March snow water equivalent in (A) UA SWE observations (1982-2017) (Broxton et al, 2019), (B) CERA-20C reanalysis (1901-2010) (Laloyaux et al, 2018), and (C) the CCSM4 Last-Millennium simulation and historical extension (850-2005) (Landrum et al, 2013). These patterns represent between 76% and 90% of the variance in their respective spatiotemporal fields.

Pacific. EOF4 is a domain-wide mode associated with temperature anomalies and SST and geopotential height anomalies off the Pacific coast and in the tropics. Although the SST correlations exhibit spatial structure resembling ENSO and other modes of Pacific SST variability (Figure 4b), none of these are significant during the 1982-2010 period (although the horseshoe-shaped PC3-SST and coastal PC4-SST patterns are significant in SST observations that extend to 2017 (Huang et al, 2017a)).

Higher order PCs/EOFs beyond the leading four also show spatially coherent variability. While these PC/EOF pairs may resemble physical climate patterns, they are not interpreted here as the orthogonality constraints may lead to mixed or otherwise poorly resolved patterns spread across multiple PCs. Given the small sample size, it can be difficult to distinguish such "degenerate multiplets" from proper modes (North et al, 1982). While the first two observed PCs are distinct modes of variability, PCs 3-4 and 5-10 are degenerate multiplets that cannot be readily distinguished from one another given

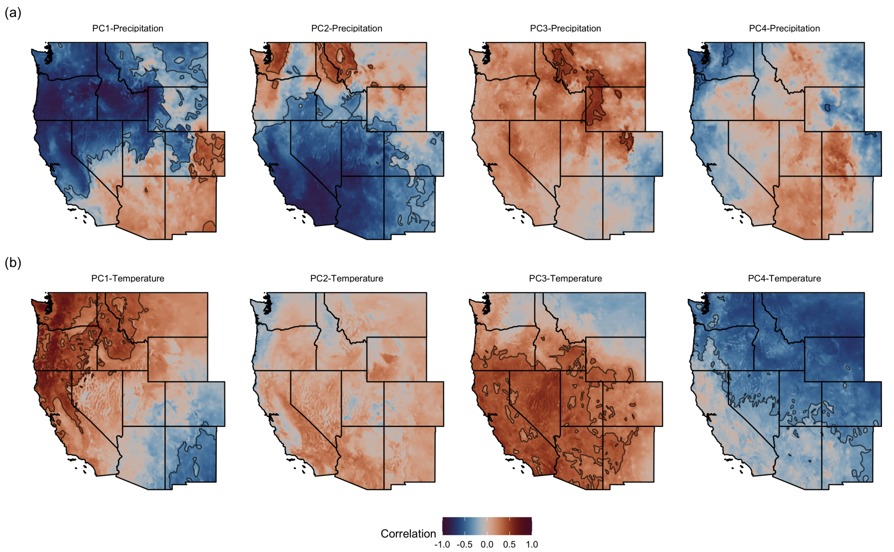


**Fig. 3** Pearson correlation coefficients between the leading four observed PC time series and October–March (A) total precipitation and (B) average temperature from PRISM (Daly et al, 2008) over the 1982-2017 period. Contour lines indicate regions of statistically significant correlation with a false discovery rate below 0.1.

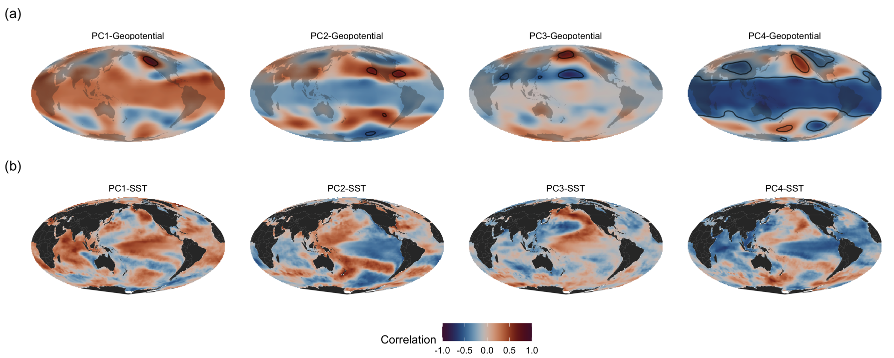


**Fig. 4** Pearson correlation coefficients between the leading four observed PC time series and October–March (A) 500mb geopotential height and (B) sea surface temperature from CERA-20C (Laloyaux et al, 2018) over the 1982-2010 period. Contour lines indicate regions of statistically significant correlation with a false discovery rate below 0.1.

the limited 36-year observational period (1982-2017). Likewise, the first four reanalysis PCs represent distinct modes while PCs 5-7 and 8-10 are degenerate multiplets. A varimax rotation of the leading ten PCs alleviates some of these concerns, yielding more discrete zones reflecting topographic interception of different directions of atmospheric flow. Regardless, that these patterns are present in reanalysis and simulation data from much longer time spans (1901-

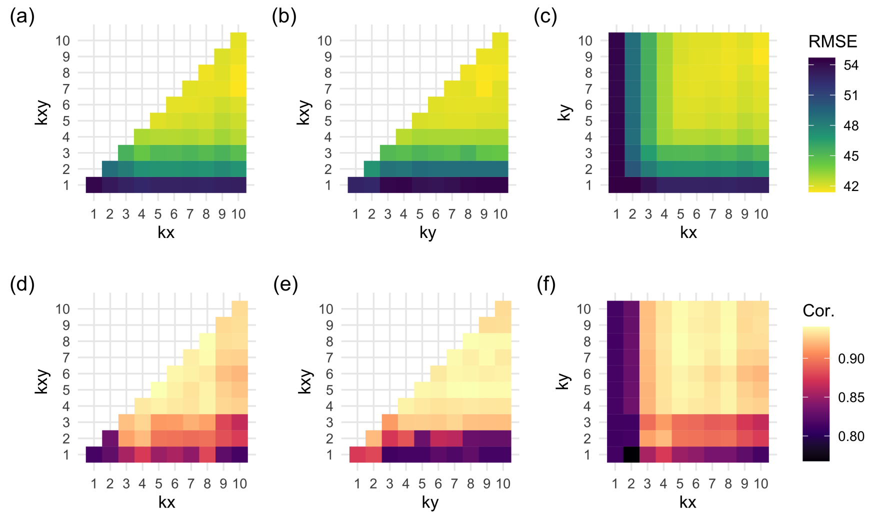


**Fig. 5** Cross validation results for the three CCA parameters after Smerdon et al (2010): the number of predictor PCs $k_x$, the number of predictand PCs $k_y$, and the number of coupled patterns $k_{xy}$. (A)–(C) cross validation results for space-time root mean square error in millimeters (lower is better). (D)–(E) correlation between observed and downscaled total domain SWE (higher is better).

2010 and 840-2005, respectively) suggests the observed patterns are robust in time and can be used as anchoring points for a non-local downscaling approach.

*Downscaling with coupled patterns has higher cross-validated skill than similar local and non-local methods.* CCA is the best-performing downscaling model under cross validation, with the lowest space-time root mean square error and effectively tied for the highest correlation with total western US SWE (Table 1). The most important parameter for model skill is the number of coupled patterns $k_{xy}$, while the precise number of prefiltering patterns $k_x$ and $k_y$ is less important as long as they are greater than or equal to the optimal number of coupled patterns (Figure 5). A CCA model with five coupled patterns maximizes the domain-wide correlation, but even one coupled pattern yields a high correlation coefficient. Likewise, a model with seven coupled patterns is the most accurate in reconstructing the entire spatiotemporal field (lowest cross validated RMSE), but a five-pattern model also performs reasonably well.

All models have comparable skill to CCA for domain-wide SWE correlations, yielding a cross validated correlation of around 0.9, but there is greater spread for space-time RMSE. Both PCR models perform similarly to CCA for domain-wide SWE correlation, but the spatial skill is degraded due to the asymmetrical relationships between the predictors and predictands. PCR and PCR-GAM models produced largely similar reconstructions, yet the nonlinear

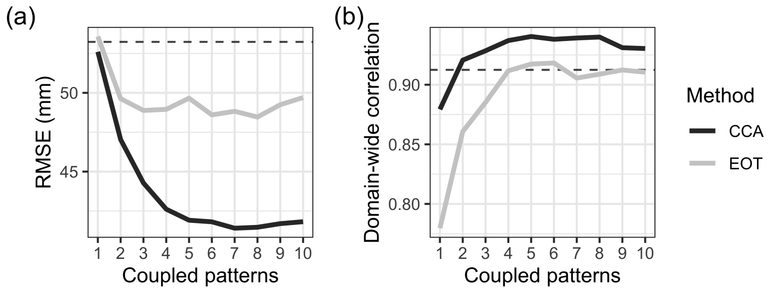


**Fig. 6** Comparison of CCA and EOT downscaling under five-fold cross validation. (A) Space-time root mean square error, in millimeters of SWE, for increasing number of coupled patterns. Lower RMSE corresponds to more accurate reconstructions. (B) Correlation between observed and reconstructed total SWE over western North America. The dashed horizontal line indicates the cross validated skill of the additive delta change model, a "local" interpolation-based downscaling approach. The curves for the PCR and PCR-GAM models (not shown) resemble those of the CCA model.

PCR-GAM consistently performs slightly worse than the linear PCR method due to its potential to overfit.

EOT yielded spatial patterns similar to the coupled-pattern methods but with notably more instability under cross validation than the pattern methods because the base grid cell tended to vary between folds (Figure 6). All methods are better than the delta change approach with additive anomalies, which performed similar to the pattern-based methods with only one or two patterns. The multiplicative delta change approach was by far the least effective, as the use of multiplicative anomalies introduced artifacts in years with unusually high SWE over areas with SWE averages close to zero. These artifacts significantly degraded the overall temporal and spatial skill, and were particularly severe under cross validation. These results support the interpretation that anchoring downscaling relationships in spatial patterns, rather than grid-cell level relationships, increases the robustness of the resulting downscaled predictions.

*Downscaling reduces spatial and temporal biases in simulated snowpack.* Downscaling the CERA-20C reanalysis with any of the above pattern-based methods considerably reduces spatial and temporal biases in the raw reanalysis. Without downscaling, the CERA-20C reanalysis tends to underpredict domain-wide total SWE averages and overpredict their variance. CCA downscaling with five coupled patterns reduces this mean and variance bias relative to observations (Figure 7). By construction, pattern-based downscaling also improves the spatial structure of simulated SWE anomalies and removes spatial biases caused by the coarse resolution of the simulated topography in a way that interpolating the simulated anomalies does not (Figure 8).

The spatial skill of the best-performing CCA model does not appear to be sensitive to recent warming trends. Using a model fit on the 80% coolest

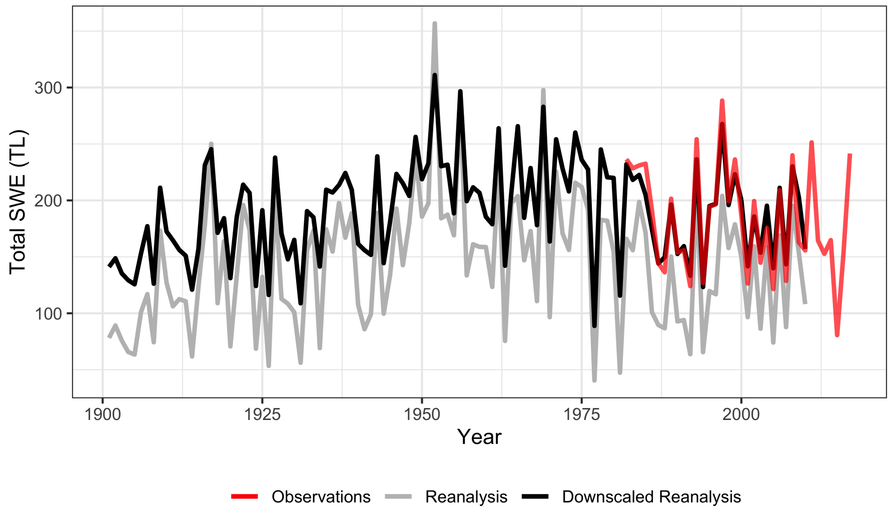


**Fig. 7** Total Western US March SWE in teraliters ($km^3$) from the CERA-20C reanalysis with five-pattern CCA downscaling (black) and without (gray), compared to recent observations (red). Downscaling adds value to the raw reanalysis by increasing the mean and decreasing the variance relative to observations.

years to predict the 20% warmest years in the calibration period (1992, 1999, 2000, 20003, 2004, 2005) yields a space-time RMSE of 40.9mm, with virtually no spatial bias between the performance of this "cool" model and the full one. However, both models do tend to underestimate the total domain SWE deficits in the driest years, suggesting that while the pattern-based methods can represent recent warming trends in space, they may still be inheriting small temporal biases from the underlying reanalysis.

Reconstructions driven instead by the NOAA-CIRES-DOE 20th century reanalysis are consistent with those downscaled from CERA-20C. The raw CERA-20C and 20CRv3 SWE fields have a domain-wide SWE correlation of 0.90 and a space-time RMSE of 77mm, while the downscaled fields have a correlation of 0.88 and RMSE of 31mm, indicating that downscaling substantially improves the spatial coherence of the reanalysis data while leaving temporal coherence largely the same. Notably, the RMSE among the two downscaled reanalysis fields is well bellow that of the best performing downscaling model under cross validation, suggesting that uncertainty due to changing calibration windows is greater than that from the selection of the particular predictor dataset.

A CCA model fit to the reanalysis data also reduces biases in the free-running CCSM4 Last Millennium simulation. Downscaling CCSM4 outputs by simply projecting them onto the patterns estimated from CERA-20C corrects mean and variance biases in total domain-wide SWE relative to the raw simulation (Figure 9). Domain-wide SWE downscaled from CCSM4 exhibits

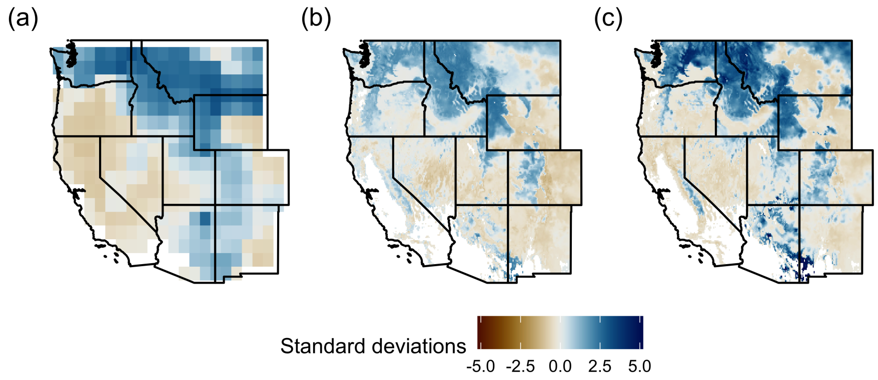


**Fig. 8** Standardized SWE anomalies for the 1997 El Niño in (A) CERA-20C reanalysis, (B) downscaled CERA-20C reanalysis using CCA with seven coupled patterns, and (C) observations. (B) and (C) are both scaled by the observed SWE standard deviation to allow comparison. Note the lower standardized anomalies in the downscaled reanalysis relative to the observations, due to the residual variance unexplained by the leading patterns.

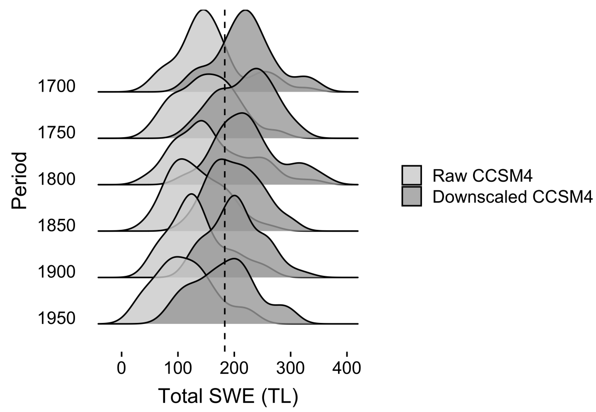


**Fig. 9** Total Western US March SWE in teraliters ($km^3$) from the CCSM4 Last Millennium simulation (Landrum et al, 2013) with CCA downscaling (dark gray) and without (light gray), compared to the 1982-2017 observed mean (dashed line). Unlike CERA-20C, CCSM4 is not constrained to be synchronous with observations and is instead assessed on a 50-year distributional basis. The same model fit from Figure 7 is used here, with the CCSM4 data simply projected onto the reanalysis PC space to enable downscaling. This approach was less successful when applied to CESM-LME outputs (not shown), as its $\sim 2°$native resolution was too coarse to meaningfully project onto the $1°$reanalysis patterns.

the same broad temporal correlations to simulated temperature and precipitation trends internal to the raw CCSM4 simulation, indicating that downscaling does not break the physical consistency of the water balance from the free-running simulation.

That simply projecting the CCSM4 data onto the CERA-20C patterns, without additional transformations, results in reasonable estimates at all is informative. For a statistical model fit on large-scale patterns from one simulation to meaningfully generalize to those from a different simulation is not guaranteed. Indeed, this is not the case for the coarser 2°CESM-LME simulation. Although the spatial patterns from CESM-LME are visually similar to those in Figure 2, they are too different at the grid cell level to be used directly for downscaling. This constraint holds regardless of whether the CESM-LME data are first resampled to the 1°resolution of CERA-20C and CCSM4 or when CERA-20C is resampled to the lower CESM resolution. This indicates that the problem is not due grid-size *per se*, but rather the impact of the simulation's native resolution on the underlying dynamics. That the downscaling model can generalize to both a distinct reanalysis dataset (20CRv3) and free-running climate model (CCSM4) at the same native resolution as the CERA-20C data used to fit the model, but not to the coarser CESM data, suggests the model generalizes well only to simulations run with a similar native resolution to the training data.

## 5 Discussion

A small number of climate modes explain the majority of observed and simulated interannual variance in snowpack across the western United States. Five to seven of these coupled modes are sufficient to downscale accurate high-resolution maps of regional snow water equivalent from coarse-resolution climate simulations. Even an extremely simple model with only one mode is able to reproduce the time evolution of the total volume of water stored in snow across the whole domain, although this is unlikely to be sufficient for full field spatiotemporal analyses. In spite of known biases in simulated SWE arising from issues of scale and process uncertainty, these findings suggest modern numerical simulations capture enough of the large scale atmosphere-ocean dynamics that drive interannual snowpack variability to be appropriate predictors for high resolution downscaling products.

Given judicious choice of physically meaningful patterns as predictors and predictands, even a simple linear downscaling method yields skillful hindcasts of observed SWE variability. This approach relies on the ability of climate and weather models to accurately simulate large-scale atmosphere-ocean variability. Rather than deriving complex transfer functions between a variety of local variables—a process that often breaks the physical consistency of climate model outputs—this approach uses the internal physical consistency of those simulations to its advantage by finding a simple mapping between simulated and observed patterns. Anchoring statistical downscaling methods in a mechanistic understanding of the climate system, instead of using downscaling as a replacement for that understanding, is of paramount importance to any downscaling project.

The leading two principal modes of variability highlighted in this study—a coherent domain-wide signal and a north/south dipole—have been identified previously in observational data of snow and several variables (Redmond and Koch, 1991; Cayan, 1996; McCabe and Dettinger, 2002; Jin et al, 2006; McCabe et al, 2013; Pederson et al, 2013; Malevich and Woodhouse, 2017). The first mode represents a domain-wide temperature anomaly associated with PNA-type atmospheric circulation. The second represents the influence of tropical Pacific SST variability (ENSO, PDO) deflecting storm tracks north or south and causing coincident temperature and precipitation anomalies in each region. This pair of influences is robust over time and appears in long-term tree-ring reconstructions from similar domains (Woodhouse, 2003; Pederson et al, 2011; Coulthard, 2015; Barandiaran et al, 2017).

There is less certainty as to the drivers of the successive modes of variability. Possible influences include cold vs. warm El Niño years, atmospheric rivers, temperature anomalies due to the Northern Annular Mode and North Atlantic Oscillation, or overlapping multidecadal modes of Pacific SST variability (QDO, PDO, IPO) (Ghatak et al, 2010; Seager et al, 2010; Barrett et al, 2015; Barandiaran et al, 2017; Goldenson et al, 2018). Complicating matters further is that the same large scale pattern can influence snowpack through multiple physical pathways and different teleconnections can act through the same pathway (Mote, 2003; Ge et al, 2009; Ghatak et al, 2010). For example, ENSO variability influences both temperature and precipitation, and by extension snow accumulation and ablation, simultaneously. Likewise, Pacific SST variability can influence storm tracks across multiple spatial and temporal scales.

Ultimately, these large-scale patterns represent the outcome of nonlinear, interacting processes that may not necessarily be well represented by linear statistical methods like PCA and CCA. What may appear to be distinct climatic modes in a PCA may instead reflect the method's linearity assumptions and orthogonality constraints. While these methods are nevertheless useful for downscaling because they isolate the parsimonious subspace of variability most influenced by these large-scale dynamics, interpretations of the individual modes must always be treated with caution. An alternative approach would be to use nonlinear feature extraction methods such as independent components analysis, self-organizing maps, or variational autoencoders to generate statistically independent patterns with increased interpretability and out-of-sample predictability (Reusch et al, 2005; Fassnacht and Derry, 2010; Henderson et al, 2017; Baño-Medina et al, 2020; He and Eastman, 2020). However, the risk of overfitting nonlinear methods remains high given the short observational window, and standard linear methods are already highly skillful.

Regardless of whether this large-scale variability is captured by linear or nonlinear methods, a degree of unexplained local variability will remain. About 20% of the local SWE variance observed at the grid cell level is left unexplained by the large-scale patterns. By definition, methods that use a restricted number of patterns on the left hand side of the regression equation will explain only a subset of the observed variance. Ideally, a downscaled SWE product would preserve this full range of variability and give some insight into the uncertainty in the downscaled estimates (Hewitt et al, 2018). An intuitive approach would be to add the residual variance back to each grid cell as uncorrelated white noise. However, we find here that the residual fraction is non-normal, spatially autocorrelated, and varies in magnitude across the study domain. While an analytical solution to the CCA noise fraction exists (Wilks, 2014), a more pragmatic approach may be to fit Gaussian process or copula models to the cross-validated errors directly. Regardless of the precise method, this residual internal variability should be modeled in order to yield downscaled data appropriate for localized climate-change impact assessments (Towler et al, 2017).

To be truly useful to researchers, stakeholders, and policy makers in the western US, downscaled snowpack products should take advantage of the wide range of long-term paleoclimate simulations to generate long-term ensembles of high-resolution snowpack variability. Such products would provide a crucial baseline for assessing present and future climate changes. Downscaled SWE estimates can also serve as spatially-explicit priors for data-assimilation (Huang et al, 2017b; Devers et al, 2019; Fiddes et al, 2019; Girotto et al, 2020), combining high-resolution snowpack fields with snow-sensitive tree-ring proxies (Coulthard et al, 2021) to generate integrated paleoclimate reconstructions (Hakim et al, 2016). We applied our reanalysis-based downscaling approach to a free-running CCSM4 simulation to test the generality of the leading SWE patterns, suggesting that pattern-based downscaling of long-term paleoclimate simulations is indeed possible. Operational downscaling for long-term climate-change impact assessments will require further steps to ensure the robustness of the coupled patterns, such as Common EOF analysis on combined reanalysis and GCM fields (Benestad, 2001) and perfect model experiments (Maraun and Widmann, 2018) to determine whether a long-term climate change signal can be captured by changes in the relative expression of existing spatial patterns. Nevertheless, our results indicate that leading modes of snowpack variability have been sufficiently stable for at least the past few centuries, and that pattern-based downscaling provides clear added value for assessing changing snowpack over the long-term.

## Declarations

### Funding

This research was supported by the National Science Foundation Paleo Perspectives on Climate Change (P2C2) grant AGS-1803995.

### Conflicts of interest/Competing interests

The authors have no relevant financial or non-financial interests to disclose.

## Availability of data and material

The daily 4km UA-SWE product can be accessed at https://nsidc.org/data/nsidc-0719 and the CERA-20C reanalysis at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/cera-20c. Although the entire analysis can be made using publicly available data, intermediate data required for the main analysis are included with the R research compendium linked below. The downscaled CERA-20C SWE estimates for 1901-2010 generated in this study will are will be made available at doi:10.5281/zenodo.5110319 upon acceptance.

## Code availability

Code for reproducing the main analysis and figures is available at https://github.com/nick-gauthier/tidyEOF and will be permanently archived on Zenodo upon acceptance at doi:10.5281/zenodo.5110395.

## References

Anderson BT, McNamara JP, Marshall HP, Flores AN (2014) Insights into the physical processes controlling correlations between snow distribution and terrain properties. Water Resources Research 50:1–19, DOI 10.1002/2013WR013714

Appelhans T, Detsch F, Nauss T (2015) Remote: Empirical orthogonal teleconnections in R. Journal of Statistical Software 65(10):1–19, DOI 10.18637/jss.v065.i10

Bales RC, Molotch NP, Painter TH, Dettinger MD, Rice R, Dozier J (2006) Mountain hydrology of the western United States. Water Resources Research 42(8):1–13, DOI 10.1029/2005WR004387

Baño-Medina J, Manzanas R, Gutierrez JM (2020) Configuration and intercomparison of deep learning neural models for statistical downscaling. Geoscientific Model Development 13(4):2109–2124, DOI 10.5194/gmd-13-2109-2020

Barandiaran D, Wang SY, De Rose RJ (2017) Gridded snow water equivalent reconstruction for Utah using forest inventory and analysis tree-ring data. Water 9(6):1–13, DOI 10.3390/w9060403

Barnett TP, Preisendorfer R (1987) Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis. Monthly Weather Review 115(9):1825–1850, DOI 10.1175/1520-0493(1987)115<1825:oaloma>2.0.co;2

Barrett BS, Henderson GR, Werling JS (2015) The Influence of the MJO on the intraseasonal variability of Northern Hemisphere spring snow depth. Journal of Climate 28(18):7250–7262, DOI 10.1175/JCLI-D-15-0092.1

Bartoń K (2020) MuMIn: Multi-model inference. R package version 1.43.17. URL https://cran.r-project.org/package=MuMIn

Benestad RE (2001) A Comparison Between Two Empirical Downscaling Strategies. International Journal of Climatology 21:1645–1668, DOI 10.1002/joc.703

Benestad RE, Chen D, Mezghani A, Fan L, Parding K (2015) On using principal components to represent stations in empirical-statistical downscaling. Tellus, Series A: Dynamic Meteorology and Oceanography 6(1), DOI 10.3402/tellusa.v67.28326

Bretherton CS, Smith C, Wallace JM (1992) An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. Journal of Climate 5(6):541–560, DOI 10.1175/1520-0442(1992)005<0541:aiomff>2.0.co;2

Broxton PD, Dawson N, Zeng X (2016a) Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. Earth and Space Science 3:246–256, DOI 10.1002/2016EA000174

Broxton PD, Zeng X, Dawson N (2016b) Why do global reanalyses and land data assimilation products underestimate snow water equivalent. Journal of Hydrometeorology 17(11):2743–2761, DOI 10.1175/JHM-D-16-0056.1

Broxton PD, Zeng X, Dawson N (2019) Daily 4 km Gridded SWE and Snow Depth from Assimilated In-Situ and Modeled Data over the Conterminous US, Version 1. DOI 10.5067/0GGPB220EX6A

Cayan DR (1996) Interannual climate variability and snowpack in the western United States. Journal of Climate 9(5):928–948, DOI 10.1175/1520-0442(1996)009<0928:ICVASI>2.0.CO;2

Chegwidden OS, Nijssen B, Rupp DE, Arnold JR, Clark MP, Hamman JJ, Kao SC, Mao Y, Mizukami N, Mote PW, Pan M, Pytlak E, Xiao M (2019) How Do Modeling Decisions Affect the Spread Among Hydrologic Climate Change Projections? Exploring a Large Ensemble of Simulations Across a Diversity of Hydroclimates. Earth's Future 7(6):623–637, DOI 10.1029/2018EF001047

Clark MP, Hendrikx J, Slater AG, Kavetski D, Anderson B, Cullen NJ, Kerr T, Örn Hreinsson E, Woods RA (2011) Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. Water Resources Research 47(7), DOI 10.1029/2011WR010745

Coulthard B (2015) Multi-century records of snow water equivalent and streamflow drought from energy-limited tree rings in south coastal British Columbia. PhD thesis, University of Victoria

Coulthard BL, Anchukaitis KJ, Pederson GT, Cook E, Littell J, Smith DJ (2021) Snowpack signals in north american tree rings. Environmental Research Letters 16(3):034,037

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28:2031–2064, DOI 10.1002/joc.1688

Devers A, Vidal J, Lauvernet C, Graff B, Vannier O (2019) A framework for high-resolution meteorological surface reanalysis through offline data assimilation in an ensemble of downscaled reconstructions. Quarterly Journal of the Royal Meteorological Society 146:153–173, DOI 10.1002/qj.3663

Dong C (2018) Remote sensing, hydrological modeling and in situ observations in snow cover research: A review. Journal of Hydrology 561(March):573–583, DOI 10.1016/j.jhydrol.2018.04.027, URL https://doi.org/10.1016/j.jhydrol.2018.04.027
van den Dool HM, Saha S, Johansson Å (2000) Empirical orthogonal teleconnections. Journal of Climate 13(8):1421–1435, DOI 10.1175/1520-0442(2000)013<1421:EOT>2.0.CO;2
Dozier J, Bair EH, Davis RE (2016) Estimating the spatial distribution of snow water equivalent in the world's mountains. WIREs Water 3(3):461–474, DOI 10.1002/wat2.1140
Erickson TA, Williams MW, Winstral A (2005) Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. Water Resources Research 41(4):1–17, DOI 10.1029/2003WR002973
Fassnacht SR, Derry JE (2010) Defining similar regions of snow in the Colorado River Basin using self-organizing maps. Water Resources Research 46(W04507), DOI 10.1029/2009WR007835
Fiddes J, Aalstad K, Westermann S (2019) Hyper-resolution ensemble-based snow reanalysis in mountain regions using clustering. Hydrology and Earth System Sciences 23(11):4717–4736, DOI 10.5194/hess-23-4717-2019
Fyfe JC, Derksen C, Mudryk L, Flato GM, Santer BD, Swart NC, Molotch NP, Zhang X, Wan H, Arora VK, Scinocca J, Jiao Y (2017) Large near-term projected snowpack loss over the western United States. Nature Communications 8(14996), DOI 10.1038/ncomms14996
Ge Y, Gong G, Frei A (2009) Physical mechanisms linking the winter Pacific-North American teleconnection pattern to spring North American snow depth. Journal of Climate 22(19):5135–5148, DOI 10.1175/2009JCLI2842.1
Ghatak D, Gong G, Frei A (2010) North American temperature, snowfall, and snow-Depth response to winter climate modes. Journal of Climate 23(9):2320–2332, DOI 10.1175/2009JCLI3050.1
Girotto M, Musselman KN, Essery RL (2020) Data Assimilation Improves Estimates of Climate-Sensitive Seasonal Snow. Current Climate Change Reports 6:81–94, DOI 10.1007/s40641-020-00159-7
Goldenson N, Leung LR, Bitz CM, Blanchard-Wrigglesworth E (2018) Influence of atmospheric rivers on mountain snowpack in the western United States. Journal of Climate 31(24):9921–9940, DOI 10.1175/JCLI-D-18-0268.1
Hakim GJ, Emile-Geay J, Steig EJ, Noone D, Anderson DM, Tardif R, Steiger N, Perkins WA (2016) The Last Millennium Climate Reanalysis Project: Framework and First Results. Journal of Geophysical Research: Atmospheres 98195:1–56, DOI 10.1002/2016JD024751, URL http://doi.wiley.com/10.1002/2016JD024751
Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: A review. International Journal of Climatology 27:1119–1152, DOI 10.1002/joc.1499

He J, Eastman JR (2020) A sequential autoencoder for teleconnection analysis. Remote Sensing 12(5):53–54, DOI 10.3390/rs12050851

Henderson GR, Barrett BS, South K (2017) Eurasian October snow water equivalent: using self-organizing maps to characterize variability and identify relationships to the MJO. International Journal of Climatology 37(2):596–606, DOI 10.1002/joc.4725

Hewitt J, Hoeting JA, Done JM, Towler E (2018) Remote effects spatial process models for modeling teleconnections. Environmetrics 29(8):1–14, DOI 10.1002/env.2523, 1612.06303

Huang B, Thorne PW, Banzon VF, Boyer T, Chepurin G, Lawrimore JH, Menne MJ, Smith TM, Vose RS, Zhang HM (2017a) NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5. DOI 10.7289/V5T72FNM

Huang C, Newman AJ, Clark MP, Wood AW, Zheng X (2017b) Evaluation of snow data assimilation using the ensemble Kalman filter for seasonal streamflow prediction in the western United States. Hydrology and Earth System Sciences 21(1):635–650, DOI 10.5194/hess-21-635-2017

Huning LS, AghaKouchak A (2020) Approaching 80 years of snow water equivalent information by merging different data streams. Scientific Data 7(1), DOI 10.1038/s41597-020-00649-1

Ikeda K, Rasmussen R, Liu C, Newman A, Chen F, Barlage M, Gutmann E, Dudhia J, Dai A, Luce C, Musselman K (2021) Snowfall and snowpack in the Western U.S. as captured by convection permitting climate simulations: current climate and pseudo global warming future climate. Climate Dynamics DOI 10/gj6v3k, URL https://link.springer.com/10.1007/s00382-021-05805-w

Jennings KS, Molotch NP (2019) The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient. Hydrology and Earth System Sciences 23:3765–3786, DOI https://doi.org/10.5194/hess-23-3765-2019

Jin J, Miller NL, Sorooshian S, Gao X (2006) Relationship between atmospheric circulation and snowpack in the western USA. Hydrological Processes 20(4):753–767, DOI 10.1002/hyp.6126

Kapnick S, Hall A (2012) Causes of recent changes in western North American snowpack. Climate Dynamics 38(9-10):1885–1899, DOI 10.1007/s00382-011-1089-y

Klos PZ, Link TE, Abatzoglou JT (2014) Extent of the rain-snow transition zone in the western U.S. under historic and projected climate. Geophysical Research Letters 41:4560–4568, DOI 10.1002/2014GL060500

Krinner G, Derksen C, Essery R, Flanner M, Hagemann S, Clark M, Hall A, Rott H, Brutel-Vuilmet C, Kim H, Ménard CB, Mudryk L, Thackeray C, Wang L, Arduini G, Balsamo G, Bartlett P, Boike J, Boone A, Chéruy F, Colin J, Cuntz M, Dai Y, Decharme B, Derry J, Ducharne A, Dutra E, Fang X, Fierz C, Ghattas J, Gusev Y, Haverd V, Kontu A, Lafaysse M, Law R, Lawrence D, Li W, Marke T, Marks D, Ménégoz M, Nasonova O, Nitta T, Niwano M, Pomeroy J, Raleigh MS, Schaedler G, Semenov

V, Smirnova TG, Stacke T, Strasser U, Svenson S, Turkov D, Wang T, Wever N, Yuan H, Zhou W, Zhu D (2018) ESM-SnowMIP: Assessing snow models and quantifying snow-related climate feedbacks. Geoscientific Model Development 11(12):5027–5049, DOI 10.5194/gmd-11-5027-2018

Laloyaux P, de Boisseson E, Balmaseda M, Bidlot JR, Broennimann S, Buizza R, Dalhgren P, Dee D, Haimberger L, Hersbach H, Kosaka Y, Martin M, Poli P, Rayner N, Rustemeier E, Schepers D (2018) CERA-20C: A Coupled Reanalysis of the Twentieth Century. Journal of Advances in Modeling Earth Systems 10(5):1172–1195, DOI 10.1029/2018MS001273

Landrum L, Otto-Bliesner BL, Wahl ER, Conley A, Lawrence PJ, Rosenbloom N, Teng H (2013) Last millennium climate and its variability in CCSM4. Journal of Climate 26(4):1085–1111, DOI 10.1175/JCLI-D-11-00326.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00326.1

Li D, Wrzesien ML, Durand M, Adam J, Lettenmaier DP (2017) How much runoff originates as snow in the western United States, and how will that change in the future? Geophysical Research Letters 44(12):6163–6172, DOI 10.1002/2017GL073551

Livezey RE, Smith TM (1999) Considerations for use of the Barnett and Preisendorfer (1987) algorithm for canonical correlation analysis of climate variations. Journal of Climate 12(1):303–305, DOI 10.1175/1520-0442-12.1.303

Malevich SB, Woodhouse CA (2017) Pacific sea surface temperatures, midlatitude atmospheric circulation, and widespread interannual anomalies in western U.S. streamflow. Geophysical Research Letters 44(10):5123–5132, DOI 10.1002/2017GL073536

Mankin JS, Diffenbaugh NS (2015) Influence of temperature and precipitation variability on near-term snow trends. Climate Dynamics 45(3-4):1099–1116, DOI 10.1007/s00382-014-2357-4, URL http://dx.doi.org/10.1007/s00382-014-2357-4

Maraun D, Widmann M (2015) The representation of location by a regional climate model in complex terrain. Hydrology and Earth System Sciences 19(8):3449–3456, DOI 10.5194/hess-19-3449-2015

Maraun D, Widmann M (2018) Statistical downscaling and bias correction for climate research. Cambridge University Press

Marks D, Dozier J (1992) Climate and energy exchange at the snow surface in the Alpine Region of the Sierra Nevada: 2. Snow cover energy balance. Water Resources Research 28(11):3043–3054, DOI 10.1029/92WR01483

Marshall AM, Abatzoglou JT, Link TE, Tennant CJ (2019) Projected Changes in Interannual Variability of Peak Snowpack Amount and Timing in the Western United States. Geophysical Research Letters 46(15):8882–8892, DOI 10.1029/2019GL083770

Maurer GE, Bowling DR (2014) Seasonal snowpack characteristics influence soil temperature and water content at multiple scales in interior western U.S. mountain ecosystems. Water Resources Research 50:5375–5377, DOI 10.1002/2013WR014452

McCabe GJ, Dettinger MD (2002) Primary modes and predictability of year-to-year snowpack variations in the Western United States from teleconnections with Pacific Ocean climate. Journal of Hydrometeorology 3(1):13–25, DOI 10.1175/1525-7541(2002)003<0013:PMAPOY>2.0.CO;2

McCabe GJ, Wolock DM (2009) Recent declines in western U.S. snowpack in the context of twentieth-century climate variability. Earth Interactions 13(12), DOI 10.1175/2009EI283.1

McCabe GJ, Betancourt JL, Pederson GT, Schwartz MD (2013) Variability common to first leaf dates and snowpack in the western conterminous United States. Earth Interactions 17(26):1–18, DOI 10.1175/2013EI000549.1

McGinnis DL (1997) Estimating climate-change impacts on colorado plateau snowpack using downscaling methods. Professional Geographer 49(1):117–125, DOI 10.1111/0033-0124.00062

Meromy L, Molotch NP, Link TE, Fassnacht SR, Rice R (2013) Subgrid variability of snow water equivalent at operational snow stations in the western USA. Hydrological Processes 27(17):2383–2400, DOI 10.1002/hyp.9355

Mote PW (2003) Trends in snow water equivalent in the Pacific Northwest and their climatic causes. Geophysical Research Letters 30(12):1–4, DOI 10.1029/2003GL017258

Mote PW, Li S, Lettenmaier DP, Xiao M, Engel R (2018) Dramatic declines in snowpack in the western US. npj Climate and Atmospheric Science 1(1):1–6, DOI 10.1038/s41612-018-0012-1, URL http://dx.doi.org/10.1038/s41612-018-0012-1

Nicholson C, Minckley TA, Shinker JJ (2019) Validating CCSM3 paleoclimate data using pollen-based reconstruction in the intermountain west. Quaternary Science Reviews 222:105,911, DOI 10.1016/j.quascirev.2019.105911, URL https://doi.org/10.1016/j.quascirev.2019.105911

North GR, Bell TL, Cahalan RF (1982) Sampling Errors in the Estimation of Empirical Orthogonal Funtions

Otto-Bliesner BL, Brady EC, Fasullo J, Jahn A, Landrum L, Stevenson S, Rosenbloom N, Mai A, Strand G (2016) Climate variability and change since 850 C.E.: An ensemble approach with the Community Earth System Model. Bulletin of the American Meteorological Society 97(5):787–801, DOI 10.1175/BAMS-D-14-00233.1

Pebesma E (2021) stars: Spatiotemporal Arrays, Raster and Vector Data Cubes. Https://r-spatial.github.io/stars/, https://github.com/r-spatial/stars/

Pederson GT, Gray ST, Woodhouse CA, Betancourt JL, Fagre DB, Littell JS, Watson E, Luckman BH, Graumlich LJ (2011) The Unusual Nature of Recent Snowpack Declines in the North American Cordillera. Science 333:332–335, DOI 10.1126/science.1201570

Pederson GT, Betancourt JL, McCabe GJ (2013) Regional patterns and proximal causes of the recent snowpack decline in the Rocky Mountains, U.S. Geophysical Research Letters 40(9):1811–1816, DOI 10.1002/grl.50424

Pierce DW, Barnett TP, Hidalgo HG, Das T, Bonfils C, Santer BD, Bala G, Dettinger MD, Cayan DR, Mirin A, Wood AW, Nozawa T (2008) Attribution of declining Western U.S. Snowpack to human effects. Journal of Climate 21(23):6425–6444, DOI 10.1175/2008JCLI2405.1

Pons MR, San-Martín D, Herrera S, Gutiérrez JM (2010) Snow trends in Northern Spain: Analysis and simulation with statistical downscaling methods. International Journal of Climatology 30(12):1795–1806, DOI 10.1002/joc.2016

R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Redmond KT, Koch RW (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, DOI 10.1029/91WR00690

Reusch DB, Alley RB, Hewitson BC (2005) Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. Polar Geography 29(3):188–212, DOI 10.1080/789610199

Rhoades AM, Ullrich PA, Zarzycki CM (2018) Projecting 21st century snowpack trends in western USA mountains using variable-resolution CESM. Climate Dynamics 50(1-2):261–288, DOI 10.1007/s00382-017-3606-0

Richman MB (1986) Review Article: Rotation of principal components. Journal of Climatology 6(March):293–335, DOI 10.1056/NEJMra1313875

Rutter N, Essery R, Pomeroy J, Altimir N, Andreadis K, Baker I, Barr A, Bartlett P, Boone A, Deng H, Douville H, Dutra E, Elder K, Ellis C, Feng X, Gelfan A, Goodbody A, Gusev Y, Gustafsson D, Hellström R, Hirabayashi Y, Hirota T, Jonas T, Koren V, Kuragina A, Lettenmaier D, Li WP, Luce C, Martin E, Nasonova O, Pumpanen J, Pyles RD, Samuelsson P, Sandells M, Schädler G, Shmakin A, Smirnova TG, Stähli M, Stöckli R, Strasser U, Su H, Suzuki K, Takata K, Tanaka K, Thompson E, Vesala T, Viterbo P, Wiltshire A, Xia K, Xue Y, Yamazaki T (2009) Evaluation of forest snow processes models (SnowMIP2). Journal of Geophysical Research Atmospheres 114(6), DOI 10.1029/2008JD011063

Schoenemann SW, Martin JT, Pederson GT, McWethy DB (2020) 2,200-Year tree-ring and lake-sediment based snowpack reconstruction for the northern Rocky Mountains highlights the historic magnitude of recent snow drought. Quaternary Science Advances 2:100,013, DOI 10.1016/j.qsa.2020.100013

Seager R, Kushnir Y, Nakamura J, Ting M, Naik N (2010) Northern Hemisphere winter snow anomalies: ENSO, NAO and the winter of 2009/10. Geophysical Research Letters 37(14), DOI 10.1029/2010GL043830

Serreze MC, Clark MP, Armstrong RL, McGinnis DA, Pulwarty RS (1999) Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. Water Resources Research 35(7):2145–2160, DOI 10.1029/1999WR900090

Siler N, Proistosescu C, Po-Chedley S (2019) Natural Variability Has Slowed the Decline in Western U.S. Snowpack Since the 1980s. Geophysical Research Letters 46(1):346–355, DOI 10.1029/2018GL081080, URL http://doi.wiley.com/10.1029/2018GL081080

Simon T, Hense A, Su B, Jiang T, Simmer C, Ohlwein C (2013) Pattern-based statistical downscaling of East Asian summer monsoon precipitation. Tellus, Series A: Dynamic Meteorology and Oceanography 65:1–12, DOI 10.3402/tellusa.v65i0.19749

Slivinski LC, Compo GP, Sardeshmukh PD, Whitaker JS, McColl C, Allan RJ, Brohan P, Yin X, Smith CA, Spencer LJ, Vose RS, Rohrer M, Conroy RP, Schuster DC, Kennedy JJ, Ashcroft L, Brönnimann S, Brunet M, Camuffo D, Cornes R, Cram TA, Domínguez-Castro F, Freeman JE, Gergis J, Hawkins E, Jones PD, Kubota H, Lee TC, Lorrey AM, Luterbacher J, Mock CJ, Przybylak RK, Pudmenzky C, Slonosky VC, Tinz B, Trewin B, Wang XL, Wilkinson C, Wood K, Wyszyński P (2020) An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3. Journal of Climate 34(4):1417–1438, DOI 10.1175/jcli-d-20-0505.1

Smerdon JE, Kaplan A, Chang D, Evans MN (2010) A pseudoproxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium. Journal of Climate 23(18):4856–4880, DOI 10.1175/2010JCLI3328.1

Tennant CJ, Harpold AA, Lohse KA, Godsey SE, Crosby BT, Larsen LG, Brooks PD, Van Kirk RW, Glenn NF (2017) Regional sensitivities of seasonal snowpack to elevation, aspect, and vegetation cover in western North America. Water Resources Research 53(8):6908–6926, DOI 10.1002/2016WR019374

Tippett MK, DelSole T, Mason SJ, Barnston AG (2008) Regression-based methods for finding coupled patterns. Journal of Climate 21(17):4384–4398, DOI 10.1175/2008JCLI2150.1

Towler E, PaiMazumder D, Holland G (2017) A framework for investigating large-scale patterns as an alternative to precipitation for downscaling to local drought. Climate Dynamics 48:881–892, DOI 10.1007/s00382-016-3116-5

Tryhorn L, Degaetano A (2013) A methodology for statistically downscaling seasonal snow cover characteristics over the Northeastern United States. International Journal of Climatology 33(12):2728–2743, DOI 10.1002/joc.3626

Van Den Dool HM (1987) A Bias in Skill in Forecasts Based on Analogues and Antilogues. Journal of Climate and Applied Meteorology 26:1278–1281

Wainer J, Cawley G (2018) Nested cross-validation when selecting classifiers is overzealous for most practical applications. arXiv preprint pp 1–9, URL http://arxiv.org/abs/1809.09446, 1809.09446

Walsh JE, Tucek DR, Peterson MR (1982) Seasonal Snow Cover and Short-Term Climatic Fluctuations over the United States. Monthly Weather Review 110(10):1474—-1486

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019) Welcome to the Tidyverse. Journal of Open Source Software 4(43):1686, DOI 10.21105/joss.01686

Wilks DS (2006) On "field significance" and the false discovery rate. Journal of Applied Meteorology and Climatology 45(9):1181–1189, DOI 10.1175/JAM2404.1

Wilks DS (2014) Probabilistic canonical correlation analysis forecasts, with application to tropical Pacific sea-surface temperatures. International Journal of Climatology 34(5):1405–1413, DOI 10.1002/joc.3771

Wilks DS (2016) "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. Bulletin of the American Meteorological Society 97(12):2263–2273, DOI 10.1175/BAMS-D-15-00267.1

Wood SN (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC press, DOI 10.1111/j.1541-0420.2006.00574.x, URL http://opus.bath.ac.uk/7011/

Woodhouse CA (2003) A 431-Yr Reconstruction of Western Colorado Snowpack from Tree Rings. Journal of Climate 16(10):1551–1561, DOI 10.2307/26249719

Xiao M, Udall B, Lettenmaier DP (2018) On the Causes of Declining Colorado River Streamflows. Water Resources Research 54(9):6739–6756, DOI 10.1029/2018WR023153

Xu Y, Jones A, Rhoades A (2019) A quantitative method to decompose SWE differences between regional climate models and reanalysis datasets. Scientific Reports 9(1):1–11, DOI 10.1038/s41598-019-52880-5

Ye K (2019) Interannual variability of March snow mass over Northern Eurasia and its relation to the concurrent and preceding surface air temperature, precipitation and atmospheric circulation. Climate Dynamics 52(5-6):2813–2836, DOI 10.1007/s00382-018-4297-x, URL http://dx.doi.org/10.1007/s00382-018-4297-x

Zeng X, Broxton P, Dawson N (2018) Snowpack Change From 1982 to 2016 Over Conterminous United States. Geophysical Research Letters 45(23):12,940–12,947, DOI 10.1029/2018GL079621