Illuminating snow droughts: The future of Western United States snowpack in the SPEAR large ensemble

Julian Francis Schmitt[1], Kai-Chih Tseng[2], Mimi Hughes[3], and Nathaniel Johnson[4]

[1]California Institute of Technology
[2]National Taiwan University
[3]NOAA Physical Sciences Laboratory
[4]Cooperative Institute for Modeling the Earth System

August 7, 2023

Abstract

Seasonal snowpack in the Western United States (WUS) is vital for meeting summer hydrological demands, reducing the intensity and frequency of wildfires, and supporting snow-tourism economies. While the frequency and severity of snow droughts (SD) are expected to increase under continued global warming, the uncertainty from internal climate variability remains challenging to quantify. Using a 30-member large ensemble from a state-of-the-art global climate model, the Seamless System for Prediction and EArth System Research (SPEAR), and an observations-based dataset, we find WUS SD changes are already significant. By 2100, SPEAR projects SDs to be nearly 9 times more frequent under shared socioeconomic pathway 5-8.5 (SSP5-8.5) and 5 times more frequent under SSP2-4.5. By investigating the influence of the two primary drivers of SD, temperature and precipitation amount, we find the average WUS SD will become warmer and wetter. To assess how these changes affect future summer water availability, we track April 15th snowpack across WUS watersheds, finding differences in the onset time of a "no-snow" threshold between regions and large internal variability within the ensemble that are both on the order of decades. For example, under SSP5-8.5, SPEAR projects California could experience no-snow anywhere between 2058 and 2096, while in the Pacific Northwest, the earliest transition happens in 2091. We attribute the inter-regional uncertainty to differences in the regions' mean winter temperature and the intra-regional uncertainty to irreducible internal climate variability. This analysis indicates that internal climate variability will remain a significant source of uncertainty for WUS hydrology through 2100.

Illuminating snow droughts: The future of Western United States snowpack in the SPEAR large ensemble

Julian Schmitt[1], **Kai-Chih Tseng**[2,3,4], **Mimi Hughes**[5], **Nathaniel C. Johnson**[2]

[1]California Institute of Technology
[2]Geophysical Fluid Dynamics Laboratory
[3]Princeton University
[4]Department of Atmospheric Sciences, National Taiwan University, 10617, Taipei, Taiwan
[5]Earth System Research Laboratories/ Physical Sciences Laboratory

Key Points:

- Severe snow droughts in the Western U.S. have increased in frequency by 26-70% across all major watersheds over the last 60 years.
- The SPEAR climate model accurately simulates the increase of Western U.S. severe snow drought that began in the early 2000s.
- SPEAR projects that increasing temperatures will cause much of the West to transition to a no-snow environment by 2100.

Corresponding author: Julian Schmitt, `jschmitt@caltech.edu`

Abstract

Seasonal snowpack in the Western United States (WUS) is vital for meeting summer hydrological demands, reducing the intensity and frequency of wildfires, and supporting snow-tourism economies. While the frequency and severity of snow droughts (SD) are expected to increase under continued global warming, the uncertainty from internal climate variability remains challenging to quantify. Using a 30-member large ensemble from a state-of-the-art global climate model, the Seamless System for Prediction and EArth System Research (SPEAR), and an observations-based dataset, we find WUS SD changes are already significant. By 2100, SPEAR projects SDs to be nearly 9 times more frequent under shared socioeconomic pathway 5-8.5 (SSP5-8.5) and 5 times more frequent under SSP2-4.5. By investigating the influence of the two primary drivers of SD, temperature and precipitation amount, we find the average WUS SD will become warmer and wetter. To assess how these changes affect future summer water availability, we track April 15th snowpack across WUS watersheds, finding differences in the onset time of a "no-snow" threshold between regions and large internal variability within the ensemble that are both on the order of decades. For example, under SSP5-8.5, SPEAR projects California could experience no-snow anywhere between 2058 and 2096, while in the Pacific Northwest, the earliest transition happens in 2091. We attribute the inter-regional uncertainty to differences in the regions' mean winter temperature and the intra-regional uncertainty to irreducible internal climate variability. This analysis indicates that internal climate variability will remain a significant source of uncertainty for WUS hydrology through 2100.

Plain Language Summary

Snow drought occurs when there is significantly less snow on the ground than normal. Snow droughts can intensify water shortages, accelerate wildfires, and harm snow-based tourism economies. For the Western United States, whose water supply is already limited, a recent increase in snow drought frequency is particularly concerning. Here, we use observational data and a new climate model to examine snow drought changes across the region between 1921 and 2100. We find snow droughts are already more common and could increase almost nine times under a business-as-usual scenario or five times under moderate emissions cuts by 2100. To better understand the increase, we tracked the evolution of the two main snow drought drivers: warmer temperatures and decreased precipitation. We find the average snow drought will become warmer and wetter, indicating warming temperatures are driving the increase. As the model consists of multiple simulations of future climate, or ensemble members, that differ only in the realization of chaotic climate variability, we can determine when Western regions are expected to lose most of their spring snowpack. We find that loss timing varies dramatically between regions and ensemble members, suggesting chaotic climate variability will shape the West's future water availability.

1 Introduction

Mountains play an indispensable role in Western United States (WUS) water supply, as their low temperatures and high precipitation capture significant water reserves in the form of snowpack. Often referred to as the "water towers" of the West, mountains store enormous amounts of winter precipitation which is measured as snow-water equivalent (SWE), or the depth of water if all snow melted instantaneously. During the dry spring and summer, the SWE is released as meltwater and supplies human populations whose water needs continue to rise (Bonsal et al., 2020). A reliable snowpack provides security to human populations across the WUS by providing water for increasing agricultural demands (Barnett et al., 2005), reducing the severity and intensity of wildfires (Trujillo et al., 2012; Gergel et al., 2017), and improving snow tourism economics (Wobus et al., 2017). According to Wobus et al. (2017), ski resorts are expected to lose 50% of ski season length by 2050 and 80% by 2090. Despite large seasonal variability, climate change has already been found to have significantly decreased SWE globally and across the WUS, particularly in late winter (Barnett et al., 2005; Kapnick & Hall, 2012; Fontrodona Bach et al., 2018; Huning & AghaKouchak, 2020). When SWE is abnormally low, the region is said to experience a snow drought (SD). SDs are driven by either warming, as a phase change from frozen to liquid, or reduced precipitation amounts. They affect the WUS's economy and human activity, even in areas far from mountain snowpack that rely on spring and summer melt waters for crop production and human consumption.

The adverse effects of SDs on a region's hydrology vary depending on the type of SD. Dry SDs, characterized by low precipitation and near- or below-normal temperatures, result in low streamflow throughout the melt season. In contrast, warm SDs occur under near- or above-normal precipitation and warm temperatures and often lead to early season snowmelt, increased spring flood risk, and summer hydrological drought (Harpold et al., 2017). While deviations from normal temperature and precipitation dictate SD occurrence, their absolute conditions impact how SDs are expected to respond to climate change. Shrestha et al. (2021) demonstrate that additional warming above a critical average winter temperature threshold of -6 to -5°C decreases snowpack. As all WUS large hydrologic unit code (HUC2) regions have historical average winter temperatures at or above -5°C, we expect their snowpack to be vulnerable to any level of warming.

To study SD across the WUS, we focus on comparing changes in SWE. Large observational uncertainty in WUS SWE measurements implies high biases are likely between any two datasets or models (Wrzesien et al., 2019). Observational model bias is driven by low sampling rates and terrain complexity, present in mountain regions, and is further magnified by assumptions in models used to generate SWE estimates (Wrzesien et al., 2019). Coupled global climate models (GCMs) are expected to produce snowpack estimates that are biased compared to observations because they have a lower spatial resolution and have temperature and precipitation biases (McCrary et al., 2017; Wrzesien et al., 2019; Kim et al., 2021; McCrary et al., 2022). Despite these biases, Matiu and Hanzer (2022) show that many models exhibit uniformity in simulating robust decreases in WUS SWE. Huning and AghaKouchak (2020), for example, have shown that SD total duration, average duration, and intensity in the WUS have increased by 28% between 1980 and 2018, and Shrestha et al. (2021) adds that these conditions are expected to continue to worsen because of the WUS's low latitude. These previous results imply that although GCMs are typically biased in their SWE base state, changes relative to their base states are still informative. As a result, we will primarily focus on comparing changes in SWE across data sets.

To investigate historical and future changes in SD frequency and intensity we use 30-member initial condition large ensembles from a state-of-the-art coupled global climate model, called the Seamless System for Prediction and EArth System Research (hereafter SPEAR) (Delworth et al., 2020). To assess SD intensity relative to the historical period, we focus on SPEAR's simulation of severe to exceptional snow droughts (D2+ SD) and follow the classification framework used by the US Drought Monitor (Svoboda et al., 2002). We first show that SPEAR accurately simulates changes in WUS SD by comparing it to an observationally based dataset and with previous studies across the historical period (1921-2011) (Livneh et al., 2013; Huning & AghaKouchak, 2020). The classifications in SPEAR show both an increase in D2+ SD occurrence across the historical period and a continued increase under future warming scenarios. To understand the conditions driving these SDs, we examine the average temperature and precipitation conditions for the study period, finding that temperature and not lack of precipitation is the main driver of the D2+ SD increase at monthly time resolution. We then provide a region-level assessment of the transition to a "no-snow" environment by the end of the 21$^{\text{st}}$ Century that accounts for scenario uncertainty and internal climate variability.

By separating the uncertainty into the portion attributable to internal climate variability and emissions uncertainty, we can determine the distribution of D2+ SD changes until 2100, the variability in the conditions that generate drought/non-drought conditions, and the probability distribution of the transition timing to a no-snow regime. We assess these changes under two scenarios in the SPEAR projections (2014-2100): a middle-of-the-road scenario (Shared Socioeconomic Pathway 2-4.5, hereafter SSP2-4.5), and a high-emissions scenario (SSP5-8.5) (Delworth et al., 2020). While the two emissions scenarios allow us to explore the effects of emissions uncertainty, the 30-member ensembles enable the estimation of internal climate variability.

2 Data and Methods

2.1 SPEAR Large Ensemble Global Climate Model

To assess changes in the probable distribution of historical and future SD, we analyzed WUS SWE in multiple 30-member SPEAR large ensembles (Delworth et al., 2020). SPEAR is a coupled global climate model recently developed at the NOAA Geophysical Fluid Dynamics Laboratory (GFDL) that is designed for improved prediction and projection on seasonal-to-multidecadal timescales. SPEAR is composed of GFDL's AM4 atmosphere, LM4 land, MOM6 ocean, and SIS2 sea-ice models. These component models are the same as GFDL's Global Climate Model version 4 (CM4) (Held et al., 2019), which is a contributor to the Coupled Model Intercomparison Project phase 6 (CMIP6). SPEAR's configuration differs from CM4 as its physical parameterization choices are optimized for climate prediction on seasonal to centennial timescales. SPEAR has a moderately high atmospheric and land-surface resolution (approximately 50 km) and a coarser ocean and sea-ice horizontal resolution of about 1°, which has meridional refinement to 0.33° at the equator. For this study, we use SPEAR's monthly SWE, temperature, and precipitation across the historical period and projections from 2014-2100 under both SSP2-4.5 and SSP5-8.5 emissions scenarios.

2.2 Observational Data

To evaluate SPEAR's historical simulation of SWE, temperature, and precipitation, we use an observations-based dataset (Livneh et al., 2013), available from 1915 to 2011, hereafter the Livneh dataset. Livneh uses statistically gridded in situ daily precipitation and temperature observations on a 1/16° grid to generate SWE estimates (among other land surface variables) using the Variable Infiltration Capacity (VIC) land model (Liang et al., 1994). To compare the Livneh dataset with the SPEAR ensemble members, we re-gridded Livneh to SPEAR's 1/2° grid and re-sampled it to SPEAR's monthly timescale. Despite incorporating observational data, gridded datasets, like Livneh, retain large uncertainties across variables including temperature, precipitation and SWE (Walton & Hall, 2018; Wrzesien et al., 2019). Many recent papers have found SWE estimates to vary widely, by upwards of a factor of 3 in some cases (Wrzesien et al., 2019), leading us to expect significant absolute biases between SWE estimates (McCrary et al., 2017, 2022). To overcome this issue, we focus our analysis on proportional changes, comparing SWE values to their own historical distributions within each dataset, and then comparing these relative changes across datasets.

We chose 1921-2011 as our historical period as it is the overlapping period of the Livneh and historical SPEAR datasets. We use the 90 complete winters to validate SPEAR and develop a baseline against which to compare the modeled future climatology. We chose to consider data at monthly resolution intervals for the following three reasons: (1) data availability, as SPEAR only recorded SWE at monthly intervals; (2) consistency with previous studies (Huning & AghaKouchak, 2020); and (3) because the monthly resolution is an appropriate timescale for monitoring snow drought.

2.3 Comparison of a Climate Large Ensemble to Observations

Delworth et al. (2020) and Maher et al. (2022) demonstrate that SPEAR accurately reproduces temperature and precipitation patterns across the US and outperforms many other state-of-the-art large ensemble climate models. Delworth et al. (2020) finds that SPEAR has negligible temperature bias and a slight positive precipitation bias across the WUS. As temperature and precipitation inform snowfall, Delworth et al. (2020) lends confidence that the underlying conditions at SPEAR's approximately 50 km resolution are well-simulated. Delworth et al. (2020), Johnson et al. (2022), and Maher et al. (2022) assess SPEAR's accuracy in representing teleconnections of the El Niño-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) to North American climate. As ENSO and PDO drive inter-annual variability across the region, assessing SPEAR's representation of these teleconnections is important for understanding how accurately the model may reproduce other extremes across the region, like SDs. Delworth et al. (2020) shows SPEAR accurately captures the relationship between PDO and North American precipitation, while Maher et al. (2022) finds that when PDO and ENSO are in phase, temperature and precipitation anomalies are amplified and vice versa. When comparing SPEAR's performance against other GCMs, Johnson et al. (2022) reports that SPEAR improves on CMIP5-generation models with a better representation of global ENSO-related temperature and precipitation patterns and Maher et al. (2022) reports SPEAR has higher accuracy and resolution than five other large ensemble models after comparing correlations of ENSO and PDO with North American winter temperature and precipitation anomalies between observations and models. Together, these studies affirm SPEAR as one of the best models to investigate changes and variability in the WUS' SWE because of its accurate representation of the response of temperature and winter hydroclimate to large-scale climate drivers.

However, as both studies focus on SPEAR's performance in reconstructing large-scale temperature and precipitation patterns, we still need to validate SWE patterns against Livneh before exploring future behavior. Livneh differs from SPEAR in that it contains only a single realization of the historical period, i.e. what actually happened, while the SPEAR ensemble captures 30 possible climates in each of its runs. The range of conditions that SPEAR's ensemble members experience is called the ensemble spread and it arises entirely from internal climate variability. Internal climate variability contributes significantly to inter-model spread in CMIP multi-model ensembles (Deser et al., 2020) and is essential for modeling extremes. When evaluating model bias, however, it means that, short of a taking a long-term average as shown in Figure 1, we do not expect biases between observations and either a single SPEAR ensemble member or the SPEAR ensemble mean to be reflective of SPEAR's accuracy in simulating the climate. While we do not expect a single SPEAR ensemble member or the ensemble mean to reproduce Livneh exactly, we do expect SPEAR to simulate a realization of the climate at least as extreme as the observed historical climate over most regions. However, with only 30 ensemble members it is still reasonable to expect an occasional observation to fall outside of the SPEAR spread. Thus, if the change in SD frequency observed in Livneh falls within the SPEAR ensemble spread, we can assume SPEAR produces a realistic historical climate. Our analysis reveals that the majority of the Livneh SWE statistics fall near clusters of SPEAR ensemble members, further strengthening the conclusion that SPEAR accurately represents the WUS climate as demonstrated in Figures 2 and S3.

2.4 Drought Classification

Before we can assess changes in SD, we first introduce our SD classification method. To ensure that only regions which typically have snow are eligible for classification, we restrict our region of study to the "historically snowy" region, areas that historically have average seasonal SWE maxima above 20 mm, based on the SPEAR ensemble mean. We then assign a classification based on how extreme each month is compared to the historical distribution of SWE across all grid cells and months.

Our methodology assigns standardized indices to each location by month and uses the US Drought Monitor's (USDM) drought classification method for hydrological drought to categorize observations into six descriptive bins: near normal (NN), abnormally dry (D0), and moderate (D1), severe (D2), extreme (D3), and exceptional (D4) drought. Wet conditions are classified analogously, with labels W0-W4 for increasingly wet months; see Figure S2 (Svoboda et al., 2002; Huning & AghaKouchak, 2020). We use a non-parametric empirical model to classify SWE, temperature, and precipitation values for each month. Without assuming the underlying distributions, a non-parametric model allows us to efficiently capture the variability without imposing subjective constraints on the data.

We begin by assigning each extended winter month of the year (Oct-April) a score based on the historical conditions at that location. Our time indices are by year (y) and month (m), e.g. $t_{1931,1}$ for January 1931, and spatial indices are in degrees latitude (i) and longitude (j). For example, $s_{40.5,250}^{t_{1931,1}}$ corresponds to a SWE value at latitude-longitude pair $(40.5, 250)$ during January 1931. We now compute an empirical distribution over $\boldsymbol{S}_{i,j}^m = \left(s_{i,j}^{t_{1921,m}}, s_{i,j}^{t_{1922,m}}, \cdots s_{i,j}^{t_{2011,m}}\right)$, representing the historical SWE values during month m at location (i, j). We then assign a value in $(0, 1)$ to each SWE measurement using the empirical cumulative distribution function, $\hat{F}_{i,j}^m$, based on the proportion of the observed data in $\boldsymbol{S}_{i,j}^m$ that fall below it. In equation 1, $\mathbb{I}(\cdot)$ takes the value 1 if SWE measurement x is larger than the historical SWE measurement, $\boldsymbol{S}_{i,j}^{t_{y,m}}$, and 0 otherwise. We sum over the historical period which ranges from 1921 to 2011, which is 91 complete years.

$$\hat{F}_{i,j}^m(x) = \frac{\text{no. of SWE values less than } x}{91} = \frac{1}{91} \sum_{y=1921}^{2011} \mathbb{I}\left(\boldsymbol{S}_{i,j}^{t_{y,m}} < x\right) \tag{1}$$

For each observed or simulated SWE value, $s_{i,j}^{t_{y,m}}$, we can then compute the z-score by plugging the SWE value into the corresponding $\hat{F}$ and then into the inverse normal distribution, Φ. We refer to these z-scores as ZSWE, which are indexed by location, month, and year. We can now classify snow droughts from the SWE value, $s_{i,j}^{t_{y,m}}$, using

$$ZSWE_{i,j}^{y,m} = \Phi\left(\hat{F}_{i,j}^m\left(s_{i,j}^{y,m}\right)\right) \tag{2}$$

Each month is then assigned a classification (W4-W0, NN, D0-D4) which can now be compared across regions. While we primarily use this framework to classify SDs, we extend the classification scheme to temperature and precipitation as needed.

A similar empirical methodology is used by Huning and AghaKouchak (2020) to classify snow droughts across the Alps, Himalayas, and WUS. Their framework is inspired by the USDM which uses the same D0-D4 classification. However, the USDM approach is not purely statistical, relying on experts to incorporate regional sensitivity into the published drought classification. Without experts, our model attempts to match the frequency of meteorological droughts in the US Drought Monitor (USDM) with snow drought frequency because the USDM is the widely accepted standard, despite its subjectivity (Svoboda et al., 2002). While our method may result in a mismatch of SWE values and impact in some locations, it provides a statistically-rigorous way to quickly capture extremes without gathering detailed human and environmental data for each pixel.

2.5 Computing Changes in Snow Drought

We can now apply our drought classification scheme to evaluate how well SPEAR reconstructs historical changes. We define two 41-year windows containing 40 complete winters to assess change, and after applying our drought classification scheme to snowpack data aggregated to the HUC2-level, we count the number of D2+ SD occurrences across the early and late historical periods, given by a ZSWE of less than −1.3, e.g. $\mathbb{I}(Z_R^t < -1.3)$ for HUC2 region R at time t. The percent change for a given region, Δ_R, is derived via

$$\Delta_R = \frac{\sum_{t'} \mathbb{I}(Z_R^{t'} < -1.3)}{\sum_t \mathbb{I}(Z_R^t < -1.3)} \cdot 100\% \text{ for } \quad t \in (1930, 1970), \quad t' \in (1971, 2011) \tag{3}$$

For example, in the Upper Colorado region, 27 months of Livneh-derived D2+ SD occur in the early historical period and 28 in the late historical period, translating to an increase of 3.7%. Next, we leverage the SPEAR ensemble spread to determine whether the overall trend is significant.

2.6 Snow Transition Threshold

In addition to evaluating drought climatology, we are also motivated to determine how a changing SWE will affect water resources. We seek to discern when a shifting climate will begin to severely and persistently impact snow as a water resource. Long-term droughts are particularly damaging, as one or two years of low snow-pack can be buffered by groundwater, above-ground reservoirs, or stored in live biomass, but these buffers dwindle with extended exposure to drought conditions. Thus, we are particularly interested in determining when no-snow conditions are expected to become systemic (Siirila-Woodburn et al., 2021; Harpold et al., 2017).

To determine this transition, we focus on April SWE because April typically corresponds to peak SWE. By first calculating the fraction of April 15$^{\text{th}}$ SWE remaining in the historically snowy portion across each of the 5 HUC2 regions: Upper Colorado, Lower Colorado, Great Basin, Pacific Northwest, and California (abbreviated UC, LC, GB, PNW, and CA), we can classify an April ($m = 4$) grid cell $s_{i,j}^{t,4}$ as no-snow for that year if there is *at most* 10% of the historical snowfall average remaining at the location (Siirila-Woodburn et al., 2021). We then calculate the regional no-snow area proportion as the fraction of the historically snowy region which experiences those conditions. Formally, we let $\mathcal{N}_R^Y$ denote this no-snow area proportion, where R represents the region, for our application a WUS HUC2, and Y the year. As before, $\bar{\boldsymbol{S}}_{i,j}^{t_Y,4}$ is the average historical SWE value for the grid cell and $s_{i,j}^{t_Y,4}$ the SWE value for the specific year. Using 10% as our no-snow threshold, $\mathbb{T} = 0.1$, then $\mathcal{N}_R^Y$ can be written as:

$$\mathcal{N}_R^Y = \frac{\sum_{(i,j)\in R} \mathbb{I}\big(s_{i,j}^{t_Y,4} < \mathbb{T} \cdot \bar{\boldsymbol{S}}_{i,j}^{4}\big)}{|(i,j) \in R|}. \tag{4}$$

Thus we have a fraction of the historically snowy region that is snow free in a given year in April. To assess when no-snow conditions become endemic, we apply a 10-year moving-window mean and then define the no-snow transition time as the year when the moving-window mean *last* crosses the area threshold, $\mathcal{A}$, before 2100. Applying this procedure to all ensemble members, we compute a distribution for when these conditions are likely to become endemic. Formally, the no-snow transition time for an ensemble member, $\mathcal{T}$, is given by:

$$\mathcal{T} := \Big[\min t : \tilde{\mathcal{N}}_R^{t'} \geq \mathcal{A} \;\forall\, t < t' \leq 2100\Big] \tag{5}$$

where $\tilde{\mathcal{N}}_R^{t'}$ gives the moving-window mean fraction of region R that experiences no-snow conditions at time t'. By requiring the moving-window average to be above $\mathcal{A}$ for all subsequent years (until 2100), $\mathcal{T}$ is uniquely determined. For a graphical explanation of this method, please refer to Figure S5.

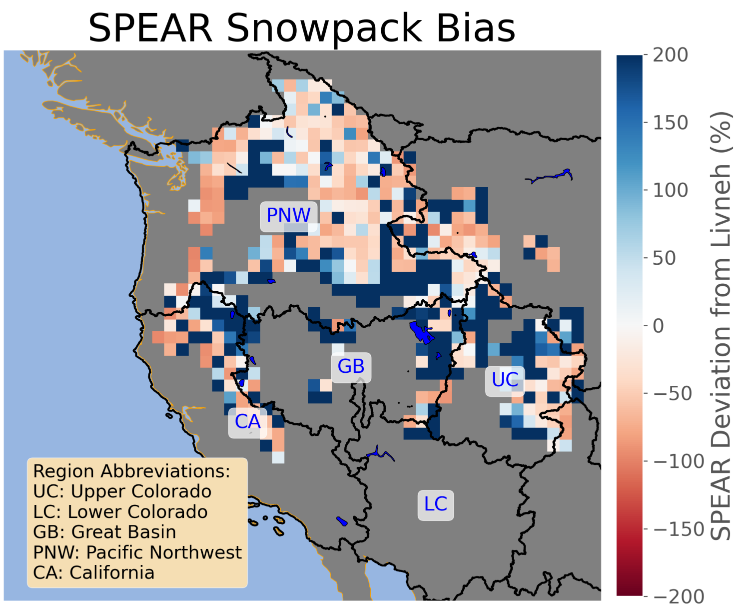

Figure 1. Winter average SPEAR SWE deviation from Livneh (%). Red indicates regions where SPEAR has a negative SWE bias while blue indicates regions with a positive bias. The five HUC2 regions are outlined in black.

3 Results

3.1 SPEAR Model Evaluation

3.1.1 SPEAR Ensemble Mean Bias

Before assessing how accurately SPEAR reconstructs historical change, we compute WUS SWE bias to assess absolute error. By taking the difference of monthly SWE averaged over the winter season (Oct-April) and the entire historical period for both datasets, we find that SPEAR has a negative snow bias across much of the Mountain West. Figure 1 reveals that in regions characterized by high elevation, SPEAR often has average SWE values less than 50% of Livneh values, while in regions adjacent to mountains, SPEAR overestimates SWE by a factor of two or more. While these are significant absolute biases, the difference is not particularly surprising because by resampling the 1/16° Livneh grid to match SPEAR's 1/2°, bias is introduced because higher elevations have disproportionately more snow than low elevations and are not accurately captured by SPEAR's 1/2° resolution due to topological smoothing (McCrary et al., 2022). We also compare historical temperature and precipitation biases in Figure S1, finding that, consistent with Delworth et al. (2020), SPEAR has a slight positive precipitation bias across the WUS.

3.1.2 Evaluating Snow Drought Changes across the Historical Period

Despite large absolute biases in SWE, SPEAR can still provide insights for future SDs if it reproduces trends and relative variability in SWE, temperature, and precipitation. Figure 2 reveals that across SPEAR, the ensemble means of all five WUS HUC2 regions experience increases in D2+ SD, ranging from an average of 26% (LC) to over 70% (UC). When we compare Livneh to the SPEAR distribution, we find that the same Livneh D2+ SD statistic always falls within the ensemble spread and is between the first and third quartiles in three of the five regions. The increases in D2+ SD occurrence are consistent with findings in Huning and AghaKouchak (2020), who use 1980-2018 as their

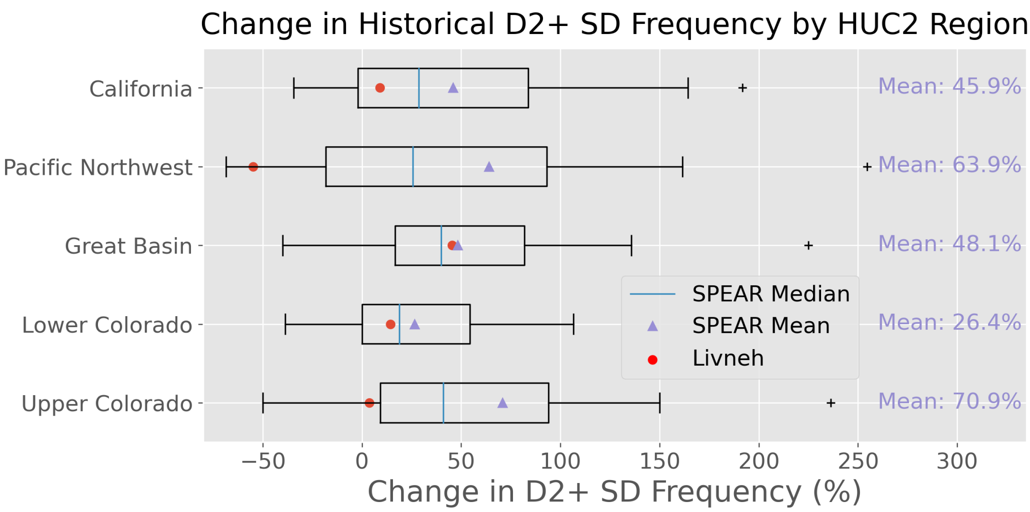

Figure 2. Comparison of SPEAR-estimated D2+ SD increases across the 1921-2011 historical period to Livneh observed increases. The SPEAR distribution is given by the box and whisker plot. The lower and upper bounds of the box correspond to the 25^{th} and 75^{th} percentiles, respectively, and points more than 1 interquartile range away from the box are denoted with a "+". The observed change in D2+ SD frequency in the Livneh dataset is marked with a red circle.

historical period — in fact, a 95% confidence interval for the SPEAR ensemble mean across four of the five regions contains the 28% benchmark for drought intensity increases found in Huning and AghaKouchak (2020), with only the UC interval exceeding the benchmark with a 30% lower bound on historical D2+ SD increases. While we could not use the same historical period due to data constraints, the agreement helps to further validate the SPEAR ensemble. See supplemental Text S1 and Figure S3 for an analysis of changes in precipitation and temperature across the historical period.

3.2 Analyzing SWE into the 21^{st} Century: Accelerating Loss

We next shift our attention to projected changes in 21^{st} century D2+ SD, focusing first on changes in droughts classified with our ZSWE metric. We construct our empirical CDF $\hat{F}^m_{i,j}$ distributions from the historical period (1921-2011) and calculate corresponding ZSWE scores for each winter month across the historically snowy west (2014-2100) for all 30 ensemble members. Projected changes in SWE are dramatic, with rapid increases in D2+ SD occurring at mid-century (Figure 3). Under SSP5-8.5, we find that towards the end of the century, all regions are projected to experience severe, extreme, or exceptional SD during most months. Under SSP2-4.5, SD increases are less severe, with conditions by the end of the century resembling conditions under SSP5-8.5 by mid-century. As expected, the higher forcing scenario corresponds with an accelerated timeline for increases in snow drought frequency. SD frequencies for all 18 study decades are shown in Figure S4.

Examining the spatial distribution of D2+ SDs in Figure 3, a pattern of regional "hot spots" emerges through time. D2+ SD frequency is consistently higher in certain regions beginning in 2030 in SSP5-8.5 and SSP2-4.5. For example, the Washington Cascades and Colorado Rockies are projected to experience more frequent D2+ SD across all decades than regions in south-central Idaho and the California Sierra Nevada. We expected to see more dramatic D2+ SD increases in the southern basins, including the Cal-

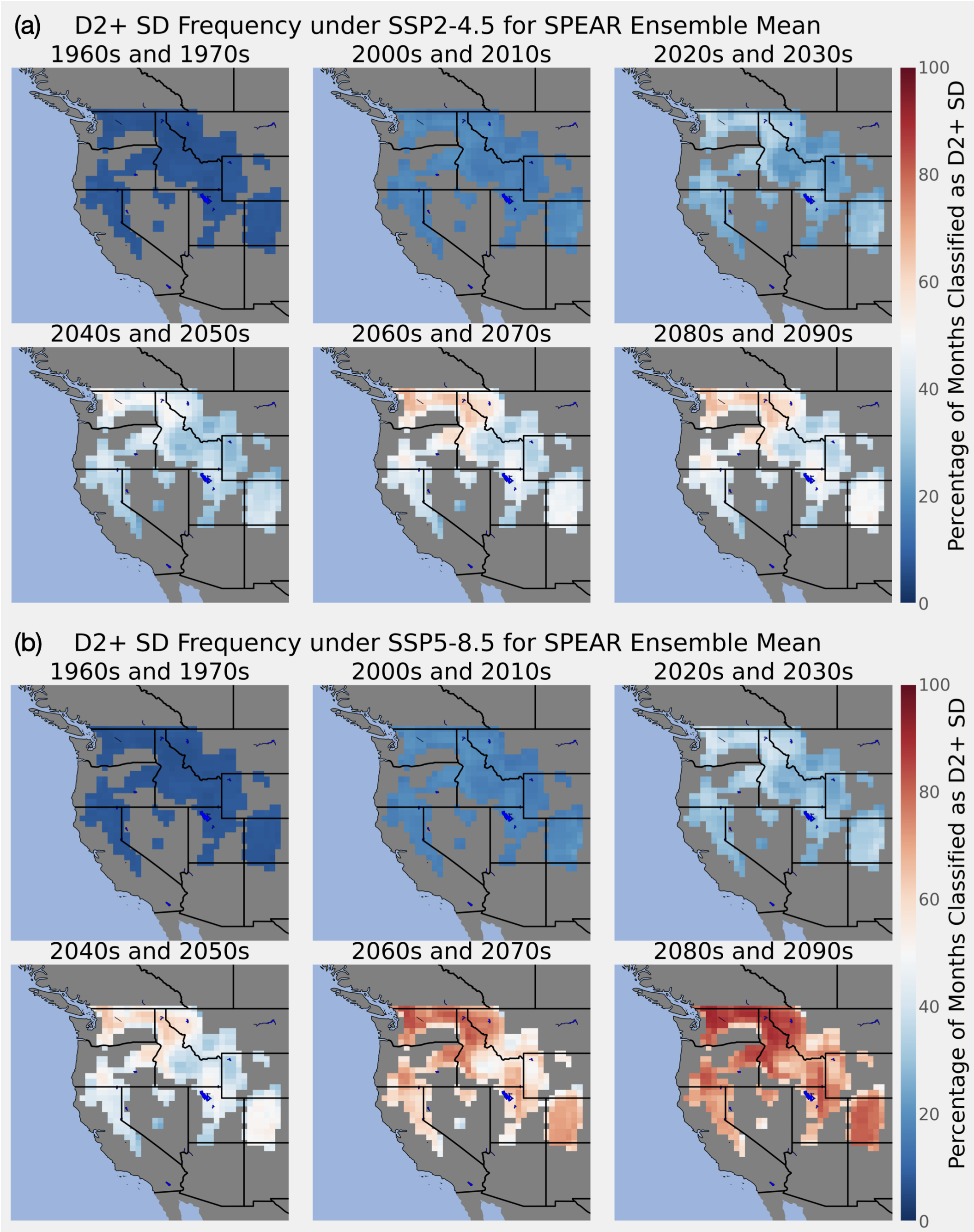

Figure 3. SPEAR D2+ SD frequencies between 1960-2100 under (a) low (SSP2-4.5) and (b) high (SSP5-8.5) emissions scenarios. The plots are masked to historically snowy regions and shaded by the percentage of winter months that the grid-cell experiences D2+ SD across a two decade period. Historically snowy regions are characterized by having an average peak SWE of at least 20mm. All 18 study decades are shown individually in Figure S4.

ifornia and Lower Colorado regions, as Shrestha et al. (2021) found that even low amounts of warming at southern latitudes result in strong SWE loss signals. We assume the hot spot pattern emerges because we are looking over a narrow enough range of latitudes that the latitude signal is overshadowed by regional variation, perhaps coming from elevation variability. Shrestha et al. (2021) examined basins ranging from the Yukon to Columbia River basins that have average winter temperatures of -8°C to +4°C, finding that below -5°C to -6°C warming temperatures did not reduce SWE. Our HUC2 regions had mean winter temperatures in historically snowy regions ranging from -5.1°C (UC) to 0.3°C (California). Therefore, we expect any amount of warming will decrease SWE and correspondingly increase D2+ SD.

While Figure 3 reveals the expected changes in D2+ SD frequency under different emissions scenarios, it does not show the impact of internal climate variability. The analysis of a large ensemble allows us to examine this effect by looking at the distribution of SD frequency across the ensemble. To study internal climate variability at the level of the entire WUS, we consider D2+ SD across the WUS in each ensemble member separately. The individual trajectories, shown in Figure 4, reveal large tail probabilities that emphasize the region may experience worse drought conditions much earlier than the ensemble mean. For example, under both future warming scenarios, the ensemble mean D2+ SD frequency is reached in some ensemble members a decade or two earlier. This emphasizes that the WUS must be prepared for D2+ SD conditions well before the ensemble mean expects them.

Figure 4 also reveals just how dramatic the increases in D2+ SD frequency may be. SPEAR ensemble members experience an average of 5-12% D2+ SD frequency during the historical period and an average of 6.5% before 2000. However, the probability of D2+ SD is projected to be over 35% by 2050 under SSP5-8.5, while under SSP2-4.5, the 35% D2+ SD probability is projected for 2070. Examining the shape of the two curves, we see an inflection point in 2000. Before 2000, both curves do not show a noticeable increase in D2+ SD frequency while after 2000 the increase is dramatic and sustained. Under SSP2-4.5, the increase in D2+ SD has a second inflection point in 2070, where the increase in snow droughts flattens. We assume the slowdown parallels the changes in the underlying climatology discussed in 3.3. Contrary to the simulations, Livneh does not show the same uptick in drought frequency in 2000. When examining the observed changes in Figure 2, we find a 53% decrease in D2+ SD frequency in the PNW. While within the SPEAR ensemble range, this decrease is far from the SPEAR ensemble mean and perhaps explains the deviation.

3.3 Temperature and Precipitation Controls on SWE

As changes in SWE are primarily driven by changes in temperature and precipitation climatology (McCrary et al., 2017; Harpold et al., 2017), we next examine changes in SWE in the phase space spanned by temperature and precipitation. By aggregating over the entire historically snowy WUS, we can determine how temperature and precipitation anomalies are driving the dramatic increase in SD. In Figure 5, each dot represents the average temperature and precipitation anomaly by decade and is colored according to the average ZSWE score. By definition, the average all-month historical (1921-2011) temperature and precipitation mean is $(0, 0)$. However, by breaking the century down by decade we can see variation within the 20$^{\text{th}}$ century.

As expected, all-month decadal averages in the historical period cluster around a zero temperature and precipitation deviation. We observe small changes in anomalies before 2000, a finding consistent with our understanding of changing D2+ SD frequency. Beginning in the 2000s, the all-month decadal-average rapidly shifts towards warmer and wetter conditions. By 2050 under SSP5-8.5, the average temperature and precipitation are 1.50 and 0.25 standard deviations higher than the 20$^{\text{th}}$ century average, respectively.

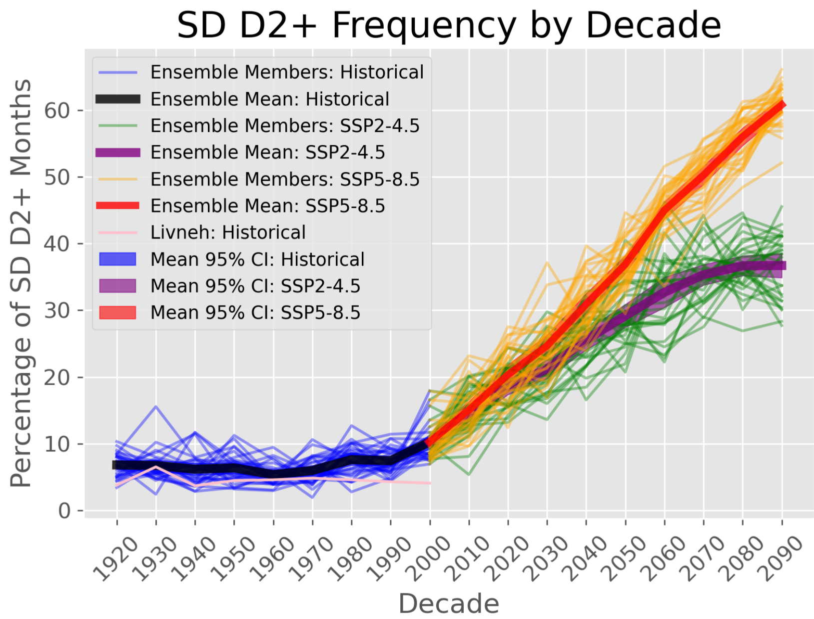

Figure 4. Each thin curve represents the percentage of historically snowy months classified as D2+ SDs and averaged by decade in Livneh (pink) and for each member of the three SPEAR ensembles; historical (blue), SSP2-4.5 (green), and SSP5-8.5 (yellow). The dark curves and surrounding shaded regions represent the ensemble mean and 95% confidence interval for the historical (blue), SSP2-4.5 (purple), and SSP5-8.5 (red) scenarios.

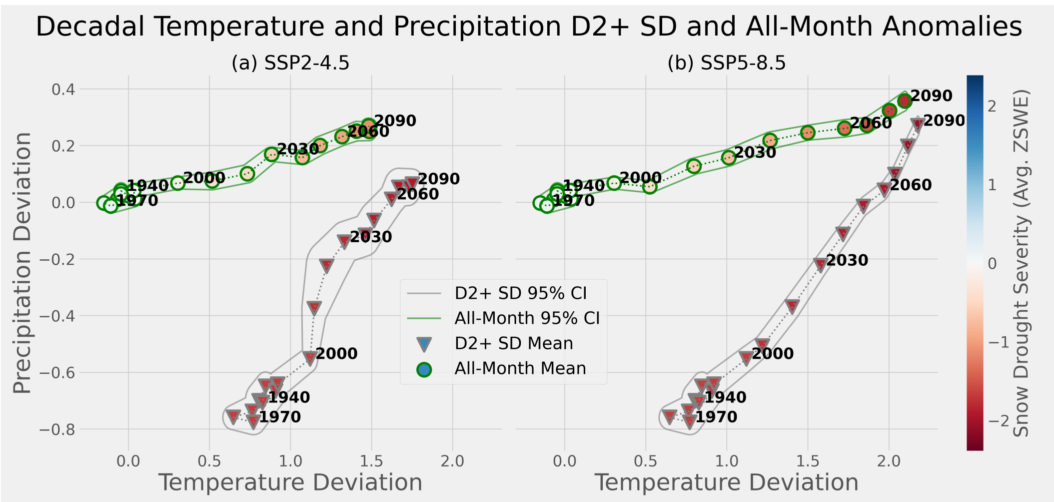

Figure 5. Temporal evolution of average temperature and precipitation anomalies with respect to the historical conditions (1921-2011). Each dot represents the average temperature and precipitation condition for historically snowy locations during winter (Oct-April) for a given decade either for all months and locations (outlined in green) or only for months classified as D2+ (outlined in gray). Each point is shaded by its average ZSWE score; thus because D2+ SD months are restricted to have a ZSWE of less than -1.3, these points average snow drought conditions are less than -1.3. Both all-month and D2+ SD-month points are surrounded by a contour which captures 95% of ensemble members. Panel (a) depicts these changes under SSP2-4.5 while (b) depicts changes under SSP5-8.5.

This corresponds to a dramatic warming and slight wetting across the WUS and indicates the average month in 2050 to be warmer than 93% of months in the historical period for a given location. For SSP2-4.5, the values are 1.18 and 0.20, respectively, reflecting a moderate increase in temperature and precipitation by mid-century, with the average month in 2050 being warmer than 88% of historical months.

To understand changes in SDs, we also track the underlying climatology of months that experience D2+ SD. Outlined in grey in Figure 5, we find historical D2+ SD averages are both dry and warm with an average temperature and precipitation anomaly of 0.6 to 0.8 and -0.6 to -0.8, respectively, indicating historical snow droughts are primarily driven by a near equal combination of both warm and dry conditions. These conditions suggest that an average historical D2+ SD month is both warmer and drier than 75% of months. However, when examining SPEAR's future climate, we find the average drought is both warmer and *wetter*. By 2050 under SSP5-8.5, the temperature deviation is 1.84 while the precipitation deviation is -0.015, indicating that future D2+ SDs are significantly warmer than the historical ones and that dry conditions are no longer needed to produce a SD. We conclude future D2+ SD conditions are driven by the increasingly high-temperature average, which is warmer than 97% of historical conditions. By 2090, the average drought month has a temperature deviation of 2.18 and a precipitation deviation of 0.27, close to the all-month anomalies of 2.10 and 0.36 for temperature and precipitation, respectively. Average monthly temperature for both D2+ and all-month averages are in the 98th percentile of historical conditions, indicating that future winter conditions will, on average, be extremely warm and that the difference between average conditions for all months and SD months has decreased. Examining the ZSWE scores for 2090 under SSP5-8.5 confirms that the convergence is also reflected in SWE changes, with the average all-month ZSWE being -1.79 and the average D2+ month having a ZSWE of -2.10. Thus, the 2090 all-month average is expected to be a D3 SD, while the average month classified as a SD is D4. Under SSP2-4.5, conditions do not reach such an extreme, with average all-month conditions by 2090 reaching 1.48 for temperature, 0.27 for precipitation, and -1.10 ZSWE. The temperature, precipitation, and ZSWE deviations for the months that experience D2+ SD are 1.75, 0.064, and -1.91, respectively. Although the gap between drought months and all-months shrinks, the difference is far less extreme than under SSP5-8.5; the average month under SSP2-4.5 is only given a D1 snow drought classification. The convergence of the all-month and drought-month temperature and precipitation anomalies, particularly under SSP5-8.5 emphasize that D2+ SDs will require increasingly smaller deviations from normal conditions to produce. This underscores that SDs will become a "new normal" for the WUS by the end of the 21st century.

3.4 Timeline for Snow-Free Conditions

In addition to changes in D2+ SD frequency, we also examine how total SWE availability is expected to change, by assessing the timing of Western regions' transition to a no-snow regime. A no-snow regime, characterized by a 10-year moving average of April SWE consistently below 10% of the historical April average, indicates severely limited summer water supply from SWE. To understand when a no-snow regime is likely to affect a HUC2 region, we examine the distribution of transition times to no-snow across SPEAR's ensemble members. By varying the area threshold, $\mathcal{A}$, we can assess how quickly conditions are expected to deteriorate. Figure 6 shows the distribution of the transition to no-snow regimes for 3 different area thresholds, $\mathcal{A}$: 50%, 75%, and 90%, for the historically snowy HUC2 regions. Note that by construction, an individual ensemble member's transition year always occurs later for higher $\mathcal{A}$. However, the ensemble distributions can overlap, which indicates large variability in the severity of conditions, especially later this century.

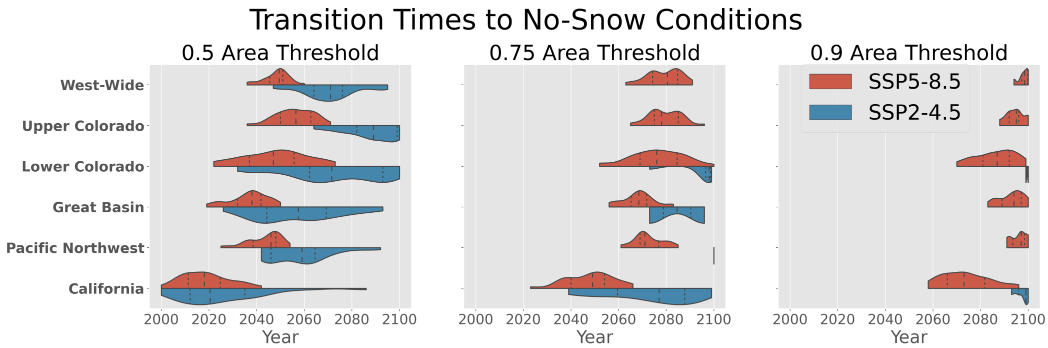

Figure 6. Distribution of SPEAR-simulated transition times to no-snow regimes, or $\mathcal{T}$, by Western HUC2 region, split between SSP5-8.5 and SSP2-4.5 scenarios. The 3 subplots represent the different thresholds $\mathcal{A} = 50\%, 75\%$ and 90%. Meeting a higher threshold corresponds with an increased proportion of the region experiencing perennial no-snow conditions, and implies more severe conditions. The vertical lines in the distributions represent the quantiles of the ensemble members that transition. We also include a transition time for the entire historically snowy WUS, labeling it "West-Wide".

When aggregated to the entire historically snowy WUS ("West-Wide"), the average transition time for $\mathcal{A} = 50\%$ is 2071 for SSP2-4.5 and 2048 for SSP5-8.5. However, when considered as separate regions, transition times for $\mathcal{A} = 50\%$ varied from as early as 2025 (CA) to 2088 (UC) under SSP2-4.5 and 2018 (CA) to 2056 (UC) for SSP5-8.5. The snow-free transition distribution center occurs later for all regions under SSP2-4.5 scenario than SSP5-8.5. However, the difference is less pronounced in regions that experience a no-snow transition earlier, such as California. We conclude that while following a lower emissions trajectory improves the probability that transitioning to a no-snow regime will occur later, large irreducible internal climate variability could result in a transition to no-snow much sooner than the ensemble mean projects.

Another notable feature of Figure 6 is the large range of transition times within each region of the 30-ensemble member transition times. We find that in some ensemble members, the earliest transition occurs over 15 years earlier than the mean transition for many regions. For example, under the SSP5-8.5 and 90% area threshold, the first ensemble member in the Lower Colorado region transitions to no-snow in 2069 while the mean transition time of the ensemble members is not until 2086. The shape of the transition time distribution under SSP2-4.5 is also more spread out than the high emissions scenario indicating larger uncertainty in the onset of no-snow conditions. The compressed timeline is a byproduct of the rapid warming accelerating the transition to no-snow because the forcing of temperature and precipitation changes happens more quickly. Thus, internal climate variability is particularly influential in SSP2-4.5 when determining no-snow transition times, while in SSP5-8.5, the accelerated radiative forcing is the dominant effect. Furthermore, while emissions reductions improve the probability that the no-snow transition will occur later in the 21$^{\text{st}}$ century, they do not guarantee a later arrival. For example, in the PNW, a quarter of the SSP2-4.5 SPEAR members transition to no-snow before the median ensemble member under SSP5-8.5. This is particularly true for regions where the transition is projected to occur earlier in the 21$^{\text{st}}$ century, likely because scenario forcing is much more similar.

To assess the probability that a region becomes snow free over the next century, we examine the fraction of ensemble members that transition to no-snow before 2100. We model the likelihood of the transition by the maximum likelihood estimator (MLE), or fraction of ensemble members that hit the transition threshold by 2100, and display these values in Table 1. By further splitting across the low and high emissions scenarios, we can model how the likelihood also changes as a function of the radiative forcing scenario. In Table 1, we see that under SSP5-8.5, $\mathcal{A} = 75\%$ is guaranteed by 2100 across all regions. The highest threshold ($\mathcal{A} = 90\%$) is guaranteed only for California, while uncertainty remains for the other 4 HUC2s. Conditions by 2100 are much less severe under SSP2-4.5, with only $\mathcal{A} = 50\%$ likely or certain for all regions, while for $\mathcal{A} = 75\%$, only California is very likely to transition to a low-snow regime; the other regions have low probability of doing so. For $\mathcal{A} = 90\%$ it is unlikely that any region will have transitioned by 2100 under SSP2-4.5.

Furthermore, when we compare the likelihood of transition to no-snow conditions with the historical regionally averaged winter temperature, we find the coldest regions are least likely to transition while the warmest are most likely. For example, under SSP5-8.5 with $\mathcal{A} = 90\%$, the order of regions by cold to warm average winter temperature and lowest to highest transition probability is the same: UC (-5.1°C, 30%), PNW (-3.9°C, 53%), GB (-2.4°C, 70%), LC (-0.7°C, 83%), and CA (0.3°C, 100%). Like Shrestha et al. (2021), we find that warming any region with a winter average temperature to greater than -5°C negatively impacts SWE. We also find that warmer regions are expected to experience a greater increase in no-snow conditions, emphasizing the role historical temperature has in determining not only whether a region will see decreased SWE but also the magnitude of the change.

Table 1 indicates that under either SSP2-4.5 or SSP5-8.5 we expect at least half of the historically snowy WUS to have less than 10% of its historical April SWE by 2100. Both columns where $\mathcal{A} = 50\%$ show greater than 80% probability for all regions, with the threshold guaranteed under SSP5-8.5. We also find that under SSP5-8.5, 4 of the 5 Western watersheds are more likely than not to cross the $\mathcal{A} = 90\%$ no-snow threshold by 2100. Upper Colorado is the exception with only a 30% chance, likely driven by lower average winter temperatures. While severe, it is important to consider how snow-covered area and total snow volume differ. As SWE declines are dominated by losses at lower elevations that are closer to the freezing point (Mote et al., 2005; Minder, 2010), we expect the topological smoothing of SPEAR may result in an overestimate of the total amount of SWE storage lost. Therefore we expect the area-based no-snow transition to over-predict the hydrological impact of warming.

4 Summary

In this study, we analyze large ensembles from a coupled global climate model, SPEAR, to understand changes in SWE across the 20th and 21st centuries. According to SPEAR, the frequency of D2+ SD has already increased dramatically across the historical period, with an average increase across all regions of 51%. While higher than the estimate of 28% in observational data found by Huning and AghaKouchak (2020), the large amount of internal climate variability of WUS SWE within the SPEAR large ensemble indicates that chaotic climate variability could account for some of the difference. SPEAR projects even more dramatic changes to come by 2100, classifying over 35% of winter months as snow droughts under RCP2-4.5 and 60% under RCP5-8.5 compared with a normalized 9.6% across the historical period. End-of-the-century projections suggest the average monthly temperature will exceed the 93rd and 97th percentiles of historical conditions under RCP2-4.5 and RCP5-8.5, respectively, and were found to be the primary driver of increased D2+ SD. To understand when future conditions will deviate significantly from 'normal,' we applied the no-snow classification defined in Siirila-Woodburn et al. (2021) to each grid cell and across years for all SPEAR ensemble members, and aggregated on the HUC2 level. We found that for the most severe threshold, $\mathcal{A} = 90\%$, a no-snow transition was more likely than not in four out of the five WUS HUC2s, the UC region being the exception. Under RCP2-4.5, only $\mathcal{A} = 50\%$ was likely for all regions. Furthermore, our

Probability of No-Snow Transition by 2100

	SSP2-4.5: Area Threshold			SSP5-8.5: Area Threshold		
HUC2 Region	50%	75%	90%	50%	75%	90%
Upper Colorado	83	0	0	100	97	30
Lower Colorado	87	23	7	100	100	83
Great Basin	100	7	0	100	100	70
Pacific Northwest	100	3	0	100	100	53
California	100	93	17	100	100	100
West-Wide	100	0	0	100	100	20

Table 1. Probability of a snow free transition occurring before 2100 at the 3 thresholds $\mathcal{A}$ based on the fraction of ensemble members who transition to a no-snow regime by 2100. We show the probabilities by area threshold, 50%, 75%, and 90%, across SSP2-4.5 and SSP5-8.5 for the historically snowy portions of each of the 5 Western HUC2 regions.

finding that California is expected to transition to no snow earlier than most regions, and Upper Colorado later, is consistent with Siirila-Woodburn et al. (2021) who use different climate models in their analysis. These conclusions emphasizes the role of future emissions in determining the no-snow transition timing.

We found regions with higher average winter temperatures were more likely to experience a transition to no-snow. The Lower Colorado and California regions, which have the highest average winter temperatures, also had the highest probability of reaching no-snow conditions across both emissions scenarios and all area thresholds. The Pacific Northwest and Upper Colorado, the regions with the coldest average temperatures, had the smallest transition probabilities. This finding parallels Shrestha et al. (2021), who found a strong correlation between average basin temperatures and the sensitivity of the region's snow to warming.

5 Remarks

By using initial condition large ensembles from a state-of-the-art GCM to study SD, we can conduct a region-wide study that accounts for both radiatively forced changes and the uncertainty attributable to internal climate variability. However, while SPEAR has higher atmospheric and land resolution than most current GCMs, its 1/2° horizontal resolution is low when compared with many mountain snowpack models (Minder, 2010), which makes it unable to resolve complex mountain topography. This limitation can result in significant warm biases and less snow (Matiu & Hanzer, 2022). We expect this may make SPEAR snowpack estimates particularly sensitive to warming, and therefore likely to overestimate increases in SD. Furthermore, Hoylman et al. (2022) asserts that using timescales longer than 30 years for drought baseline climatology, as has been done here and in the vast majority of previous literature (Svoboda et al., 2002), can result in over-estimating the drought threat in a climate that is shifting towards (in this case) a less snowy state – although they argue that the reference period should take into consideration the adaptive capability of the system in question. Further work should investigate both the sensitivity of SD estimates to GCM resolution and the effect of reference climatology choice on drought severity estimation.

Here, we have assessed changes in SD across the WUS in a GCM, focusing on validating historical changes, assessing changes to the underlying climatology, and determining when WUS regions may essentially become snow-free. For this latter objective, we developed a metric, the no-snow transition time, to track both how soon a region is expected to change and the uncertainty of this timing attributable to internal climate variability. One promising avenue for future research is to examine SD changes over smaller regions, such as HUC4s, to determine the most vulnerable locations on a sub-region scale. This would also allow further exploration of SWE's sensitivity to latitude and elevation, although at smaller watershed scales the GCM's horizontal resolution will become more problematic. Also, estimating total SWE losses and melt timing across each region would allow us to better estimate the impacts of snow droughts on the West's hydrological system. The impacts of future SDs will be felt across the entire country, both directly from the hydrological and tourism resources that consistent snowpack provides and indirectly through loss of agricultural output from summer water shortages or drifting wildfire smoke. Understanding the probable severity and timing of when these conditions are projected to become most damaging, alongside uncertainty from emissions and internal climate variability, will allow policymakers and infrastructure planners to best prepare the West for a future with less snow.

Acknowledgments

J F S was supported by the Ernest F. Hollings Undergraduate Scholarship Program. The authors would like to thank Mike Hobbins for his thoughtful expertise on an internal review, Darren Jackson for help with the HUC2 shapefiles, and Sarah Kapnick for inspiring the project early on. Finally, they thank NOAA Physical Sciences Laboratory for hosting the publicly available Livneh dataset that was used for the project.

Data availability statement: The Livneh daily CONUS near-surface gridded meteorological and derived hydrometeorological data used in the historical analysis are available from the NCEI at `doi:10.7289/V5X34VF6` (Livneh et al., 2013). The SPEAR ensemble simulation data used to assess historical and future snow drought in this study are kept at `10.5281/zenodo.7121527`, with the full publicly-available dataset available at `https://www.gfdl.noaa.gov/spear_large_ensembles/` (Delworth et al., 2020). The HUC2 shapefiles used to aggregate the climate data are kept at `10.5281/zenodo.7121527`, which are originally from the USGS watershed boundary dataset (`https://www.usgs.gov/national-hydrography/watershed-boundary-dataset`). The scripts used for data processing and statistical analysis are preserved at `10.5281/zenodo.7130302` and developed openly on GitHub at `https://github.com/Julians42/Snow_Droughts`.

References

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. Retrieved 2022-07-26, from `https://www.nature.com/articles/nature04141` (Number: 7066 Publisher: Nature Publishing Group) doi: 10.1038/nature04141

Bonsal, B., Shrestha, R. R., Dibike, Y., Peters, D. L., Spence, C., Mudryk, L., & Yang, D. (2020). Western canadian freshwater availability: current and future vulnerabilities. *Environmental Reviews*, *28*(4), 528–545. Retrieved 2022-07-26, from `https://cdnsciencepub.com/doi/10.1139/er-2020-0040` (Publisher:

NRC Research Press) doi: 10.1139/er-2020-0040

Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., Dunne, K. A., ... Zhao, M. (2020). Spear: The next generation gfdl modeling system for seasonal to multidecadal prediction and projection. *Journal of Advances in Modeling Earth Systems*, *12*(3), e2019MS001895. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001895` (e2019MS001895 2019MS001895) doi: https://doi.org/10.1029/2019MS001895

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., ... Ting, M. (2020, April). Insights from earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. Retrieved 2022-08-31, from `https://www.nature.com/articles/s41558-020-0731-2` (Number: 4 Publisher: Nature Publishing Group) doi: 10.1038/s41558-020-0731-2

Fontrodona Bach, A., van der Schrier, G., Melsen, L. A., Klein Tank, A. M. G., & Teuling, A. J. (2018). Widespread and accelerated decrease of observed mean and extreme snow depth over europe. *Geophysical Research Letters*, *45*(22), 12,312–12,319. Retrieved 2022-07-26, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2018GL079799` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL079799) doi: 10.1029/2018GL079799

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change*, *141*(2), 287–299. Retrieved from `https://doi.org/10.1007/s10584-017-1899-y` doi: 10.1007/s10584-017-1899-y

Harpold, A., Dettinger, M., & Rajagopal, S. (2017, 02). Defining snow drought and why it matters. *Eos Transactions American Geophysical Union*. doi: 10.1029/2017EO068775

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and performance of gfdl's cm4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691-3727. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001829` doi: https://doi.org/10.1029/2019MS001829

Hoylman, Z. H., Bocinsky, R. K., & Jencso, K. G. (2022). Drought assessment has been outpaced by climate change: empirical arguments for a paradigm shift. *Nature Communications*, *13*(1), 2715. Retrieved 2022-11-12, from `https://www.nature.com/articles/s41467-022-30316-5` (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41467-022-30316-5

Huning, L. S., & AghaKouchak, A. (2020). Global snow drought hot spots and characteristics. *Proceedings of the National Academy of Sciences*, *117*(33), 19753-19759. Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.1915921117` doi: 10.1073/pnas.1915921117

Johnson, N. C., Wittenberg, A. T., Rosati, A. J., Delworth, T. L., & Cooke, W. (2022). Future changes in boreal winter ENSO teleconnections in a large ensemble of high-resolution climate simulations. *Frontiers in Climate*, *4*. Retrieved 2023-07-31, from `https://www.frontiersin.org/articles/10.3389/fclim.2022.941055` doi: 10.3389/fclim.2022.941055

Kapnick, S., & Hall, A. (2012). Causes of recent changes in western north american snowpack. *Climate Dynamics*, *38*(9), 1885–1899. Retrieved 2023-06-02, from `http://link.springer.com/10.1007/s00382-011-1089-y` doi: 10.1007/s00382-011-1089-y

Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., ... Wang, S. (2021). Snow ensemble uncertainty project (seup): quantification of snow water equivalent uncertainty across north america via ensemble land surface modeling. *The Cryosphere*, *15*(2), 771–791. Retrieved from https://tc.copernicus.org/articles/15/771/2021/ doi: 10.5194/tc-15-771-2021

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*, 14415–14428. Retrieved 2023-03-28, from https://onlinelibrary.wiley.com/doi/abs/10.1029/94JD00483 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/94JD00483) doi: 10.1029/94JD00483

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., ... Lettenmaier, D. P. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous united states: Update and extensions. *Journal of Climate*, *26*(23), 9384 - 9392. Retrieved from https://journals.ametsoc.org/view/journals/clim/26/23/jcli-d-12-00508.1.xml doi: 10.1175/JCLI-D-12-00508.1

Maher, N., Kay, J. E., & Capotondi, A. (2022). Modulation of ENSO teleconnections over north america by the pacific decadal oscillation. *Environmental Research Letters*, *17*(11), 114005. Retrieved 2022-11-01, from https://dx.doi.org/10.1088/1748-9326/ac9327 (Publisher: IOP Publishing) doi: 10.1088/1748-9326/ac9327

Matiu, M., & Hanzer, F. (2022). Bias adjustment and downscaling of snow cover fraction projections from regional climate models using remote sensing for the european alps. *Hydrology and Earth System Sciences*, *26*(12), 3037–3054. Retrieved 2022-09-21, from https://hess.copernicus.org/articles/26/3037/2022/ (Publisher: Copernicus GmbH) doi: 10.5194/hess-26-3037-2022

McCrary, R. R., McGinnis, S., & Mearns, L. O. (2017). Evaluation of snow water equivalent in narccap simulations, including measures of observational uncertainty. *Journal of Hydrometeorology*, *18*(9), 2425 - 2452. Retrieved from https://journals.ametsoc.org/view/journals/hydr/18/9/jhm-d-16-0264_1.xml doi: 10.1175/JHM-D-16-0264.1

McCrary, R. R., Mearns, L. O., Hughes, M., Biner, S., & Bukovsky, M. S. (2022, February). Projections of North American snow from NA-CORDEX and their uncertainties, with a focus on model resolution. *Climatic Change*, *170*(3), 1-25. Retrieved from https://ideas.repec.org/a/spr/climat/v170y2022i3d10.1007_s10584-021-03294-8.html doi: 10.1007/s10584-021-03294-

Minder, J. R. (2010). The sensitivity of mountain snowpack accumulation to climate warming. *Journal of Climate*, *23*(10), 2634–2650. Retrieved 2022-09-21, from http://journals.ametsoc.org/doi/10.1175/2009JCLI3263.1 doi: 10.1175/2009JCLI3263.1

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). Declining mountain snowpack in western north america*. *Bulletin of the American Meteorological Society*, *86*(1), 39 - 50. Retrieved from https://journals.ametsoc.org/view/journals/bams/86/1/bams-86-1-39.xml doi: 10.1175/BAMS-86-1-39

Shrestha, R. R., Bonsal, B. R., Bonnyman, J. M., Cannon, A. J., & Najafi, M. R. (2021). Heterogeneous snowpack response and snow drought occurrence across river basins of northwestern north america under 1.0 °c to 4.0 °c global warming. *Climatic Change*, *164*(3), 40.

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., ... Kaatz, L. (2021). A low-to-no snow future and its impacts on water resources in the western united states. *Nature Reviews Earth & Environment*, *2*(11), 800–819. Retrieved from https://doi.org/10.1038/s43017-021-00219-y doi: 10.1038/s43017-021-00219-y

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... Stephens, S. (2002). The drought monitor. *Bulletin of the American Me-

teorological Society, *83*(8), 1181 - 1190. Retrieved from https://journals.ametsoc.org/view/journals/bams/83/8/1520-0477-83_8_1181.xml doi: 10.1175/1520-0477-83.8.1181

Trujillo, E., Molotch, N. P., Goulden, M. L., Kelly, A. E., & Bales, R. C. (2012). Elevation-dependent influence of snow accumulation on forest greening. *Nature Geoscience*, *5*(10), 705–709. Retrieved 2022-07-26, from https://www.nature.com/articles/ngeo1571 (Number: 10 Publisher: Nature Publishing Group) doi: 10.1038/ngeo1571

Walton, D., & Hall, A. (2018). An assessment of high-resolution gridded temperature datasets over california. *Journal of Climate*, *31*(10), 3789–3810. Retrieved 2022-11-18, from https://journals.ametsoc.org/view/journals/clim/31/10/jcli-d-17-0410.1.xml (Publisher: American Meteorological Society Section: Journal of Climate) doi: 10.1175/JCLI-D-17-0410.1

Wobus, C., Small, E. E., Hosterman, H., Mills, D., Stein, J., Rissing, M., … Martinich, J. (2017). Projected climate change impacts on skiing and snowmobiling: A case study of the united states. *Global Environmental Change*, *45*, 1–14. Retrieved 2022-09-21, from https://www.sciencedirect.com/science/article/pii/S0959378016305556 doi: 10.1016/j.gloenvcha.2017.04.006

Wrzesien, M. L., Pavelsky, T. M., Durand, M. T., Dozier, J., & Lundquist, J. D. (2019). Characterizing biases in mountain snow accumulation from global data sets. *Water Resources Research*, *55*(11), 9873-9891. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR025350 doi: https://doi.org/10.1029/2019WR025350

Supporting Information for "Illuminating snow droughts: The future of Western United States snowpack in the SPEAR large ensemble"

Julian Schmitt[1], Kai-Chih Tseng[234], Mimi Hughes[5], Nathaniel C. Johnson[2]

[1]California Institute of Technology

[2]Geophysical Fluid Dynamics Laboratory

[3]Princeton University

[4]Department of Atmospheric Sciences, National Taiwan University, 10617, Taipei, Taiwan

[5]Earth System Research Laboratories/ Physical Sciences Laboratory

Contents of this file

Introduction

We include additional text and figures to complement the discussion in the main text.

Text S1: Historical Changes in Temperature and Precipitation Extremes

To assess SPEAR's ability to represent historical temperature and precipitation extremes, we applied the methodology used in section 2.4 of the main text for snow drought classification to temperature and precipitation. We first aggregated SPEAR and Livneh to monthly time scales, taking the maximum high and minimum low daily temperature for each month. We do so because we assume these metrics are more likely to capture heat waves and cold extremes than an average of daily high and low temperatures. For example, a severe multi-day heat wave in January has the potential to melt snowpack quickly but might fail to show up on a 30-day average. To measure changes in meteorological drought conditions, we utilize the D2+ severity notation for dry extremes, and introduce W2+ for wet extremes. For temperatures, we introduce H2+ to indicate heat extremes and C2+ to indicate cold extremes; refer to Figure S2 for notation. As our analysis focuses on both monthly maximum and minimum temperatures, we append the subscripts "max" and "min" to distinguish between these conditions, respectively. Therefore, $H2+_{min}$ represents the change in frequency of warm extremes for monthly minimum temperatures. To assess whether changes are significant across the historical time period, we evaluate a 95% confidence interval for the SPEAR ensemble mean, assuming the underlying changes were distributed normally. If the interval does not contain zero change, then the forced component is significant in SPEAR. We present these calculations in Figure S3 which assesses changes to monthly meteorological drought (D2+), warm temperature extremes ($H2+_{max}$), and cold temperature extremes ($C2+_{min}$ and $H2+_{min}$). Together, these panels provide a method to validate the SPEAR ensemble against changes in extreme temperature and precipitation observations.

Across the historical period (1930-2011), we found that while SPEAR's changes in D2+ meteorological drought were not significant, several measures of temperature extremes were. Across the five HUC2 regions, we find that H2+$_{max}$ extreme heat increased on average between 59% and 73%, C2+$_{min}$ extreme cold decreased between 18-21%, and H2+$_{min}$ extreme heat increased between 41% and 60% (Figure S3(b, c, d)). These changes indicate significant warming in both maximum and minimum temperatures and both of these trends are expected to negatively impact WUS snowpack. In both SPEAR and Livneh, extreme heat events have increased in frequency while extreme cold events have decreased on average. When we assess agreement between SPEAR and Livneh, we find that all but one Livneh observation falls within the SPEAR ensemble range, suggesting SPEAR is able to accurately reproduce changes in precipitation and temperature extremes across the historical period. Examining which trends are significant in SPEAR, we find all temperature trends to be significant while only the increase in D2+ meteorological drought in the LC region is significant. The LC saw an average increase in meteorological drought of 16% in SPEAR, while in Livneh the increase in the LC was 48%. Amongst the five HUC2 regions, the LC region increase was also the most extreme increase among Livneh meteorological drought observations (Figure S3(a)). Together, these observations may indicate the LC is drying more rapidly than other regions. When examining temperature trends, the PNW stood out as it experienced the smallest changes in extreme temperatures and was the only region to observe a decrease in meteorological drought. We assume these underlying colder, wetter conditions across the latter half of the historical period explains the decrease in D2+ SD frequency over the historical period in the PNW seen in Figure 3 and perhaps the deviation in the early 2000s of Livneh D2+ SD frequency from the

SPEAR ensemble in Figure 5. While PNW falls further from the SPEAR ensemble mean, the changes are still within the SPEAR ensemble range and thus may be attributable to internal climate variability. The strong agreement between changes to historical meteorological and temperature conditions in SPEAR and Livneh further lends confidence to SPEAR's ability to capture historical trends across the WUS.

References

Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., Dunne, K. A., ... Zhao, M. (2020). Spear: The next generation gfdl modeling system for seasonal to multidecadal prediction and projection. *Journal of Advances in Modeling Earth Systems*, *12*(3), e2019MS001895. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001895` (e2019MS001895 2019MS001895) doi: https://doi.org/10.1029/2019MS001895

Huning, L. S., & AghaKouchak, A. (2020). Global snow drought hot spots and characteristics. *Proceedings of the National Academy of Sciences*, *117*(33), 19753-19759. Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.1915921117` doi: 10.1073/pnas.1915921117

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., ... Kaatz, L. (2021). A low-to-no snow future and its impacts on water resources in the western united states. *Nature Reviews Earth & Environment*, *2*(11), 800–819. Retrieved from `https://doi.org/10.1038/s43017-021-00219-y` doi: 10.1038/s43017-021-00219-y

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... Stephens, S. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181 - 1190. Retrieved from `https://journals.ametsoc.org/view/journals/bams/83/8/1520-0477-83_8_1181.xml` doi: 10.1175/1520-0477-83.8.1181

Walton, D., & Hall, A. (2018). An assessment of high-resolution gridded temperature datasets over california. *Journal of Climate*, *31*(10), 3789–3810. Retrieved 2022-11-18, from `https://journals.ametsoc.org/view/journals/clim/31/10/jcli-d`

-17-0410.1.xml (Publisher: American Meteorological Society Section: Journal of Climate) doi: 10.1175/JCLI-D-17-0410.1

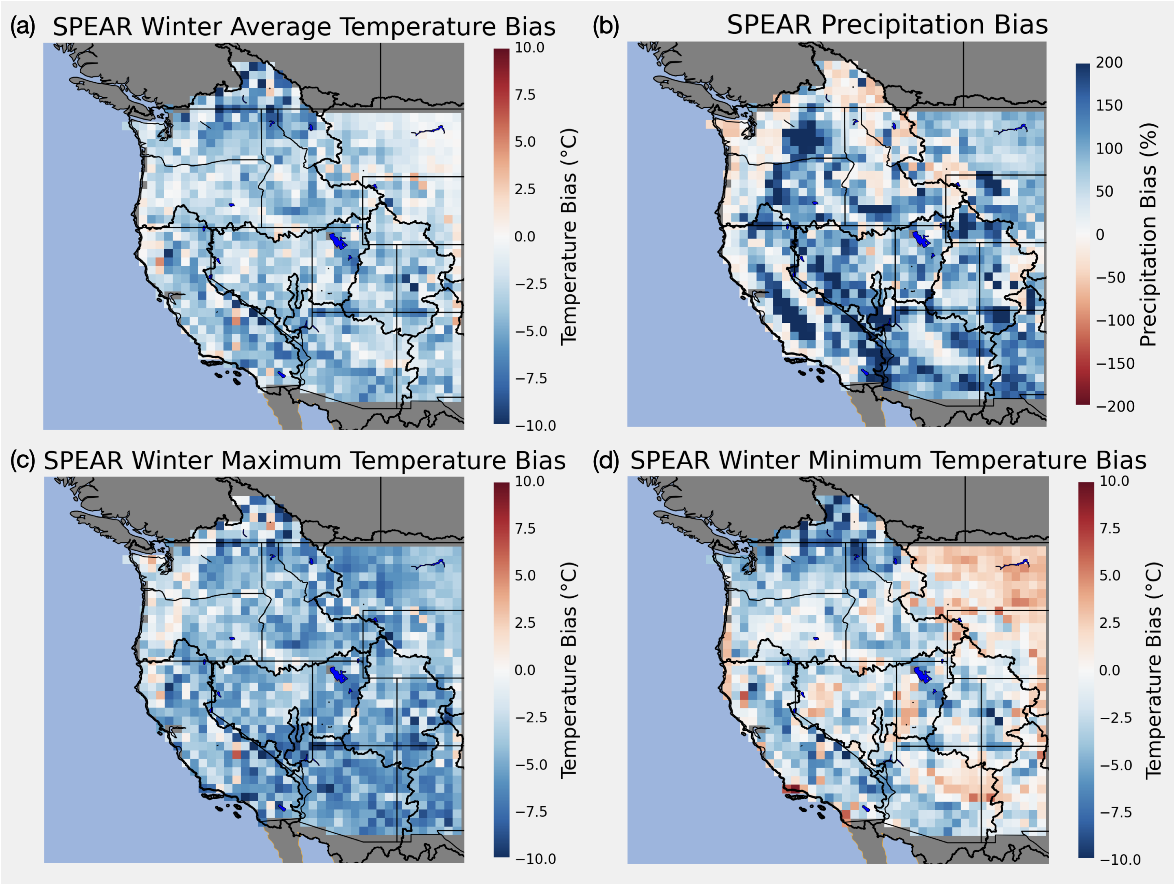

Figure S1. 90-Year winter temperature and precipitation biases including (a) average temperature bias, (b) precipitation bias, (c) maximum temperature bias, and (d) minimum temperature bias. The winter average temperature bias was computed by taking the difference of the average maximum and average minimum temperatures of SPEAR and Livneh between October 1^{st} and April 31^{st}. Overall SPEAR has a slight cold and wet bias across much of the Western United States. The wet bias is consistent with Delworth et al. (2020). By examining the maximum and minimum temperature biases we see that SPEAR has a significant cold bias for maximum temperatures across the entire WUS, while it has a systematic cold bias for minimum temperatures over mountainous regions and slight warm bias over the rest of the WUS. We expect that some of the bias can be explained by the differences in model resolution: SPEAR is on a 1/2° grid while Livneh is on a 1/16° grid. We also note that Livneh has a particularly high lapse rate of 6.5°C/1000m which may contribute some additional bias (Walton & Hall, 2018).

Classification of Extremes by Z-Score				
Drought Severity	**Temperature Severity**	**Description**	**Z-Score**	**Probability of at least as Extreme**
D4	C4	Exceptional	$Z \leq -2.0$	0.023
D3	C3	Extreme	$-2.0 < Z \leq -1.6$	0.055
D2	C2	Severe	$-1.6 < Z \leq -1.3$	0.097
D1	C1	Moderate	$-1.3 < Z \leq -0.8$	0.21
D0	C0	Abnormal	$-0.8 < Z \leq -0.5$	0.31
NN	NN	Near Normal	$-0.5 < Z < 0.5$	--
W0	H0	Abnormal	$0.5 \leq Z < 0.8$	0.31
W1	H1	Moderate	$0.8 \leq Z < 1.3$	0.21
W2	H2	Severe	$1.3 \leq Z < 1.6$	0.097
W3	H3	Extreme	$1.6 \leq Z\ 2.0$	0.055
W4	H4	Exceptional	$2 \leq Z$	0.023

Figure S2. List of drought and temperature classification abbreviations, a text description for each, the corresponding Z-Score, and the probability of an event being at least as extreme. This probability captures how likely a random historical month is to be classified in that category or one that is more extreme. This table uses identical Z-Score ranges to Huning and AghaKouchak (2020) and attempts to mimic frequencies of hydrological drought given by the US Drought Monitor (Svoboda et al., 2002).

Figure S3. Changes in wintertime (Oct-Apr) historical precipitation and temperature extremes in (a) D2+ meteorological drought and (b) H2+$_{max}$, (c) C2+$_{min}$, and (d) H2+$_{min}$ temperatures. The shaded histogram depicts the SPEAR ensemble distribution, with ensemble mean and confidence interval marked with vertical black dashed and solid lines, respectively. The observed value from the Livneh dataset is marked as a vertical line shaded red in (a) and blue in (b-d). A vertical dotted zero trend line is included for reference.

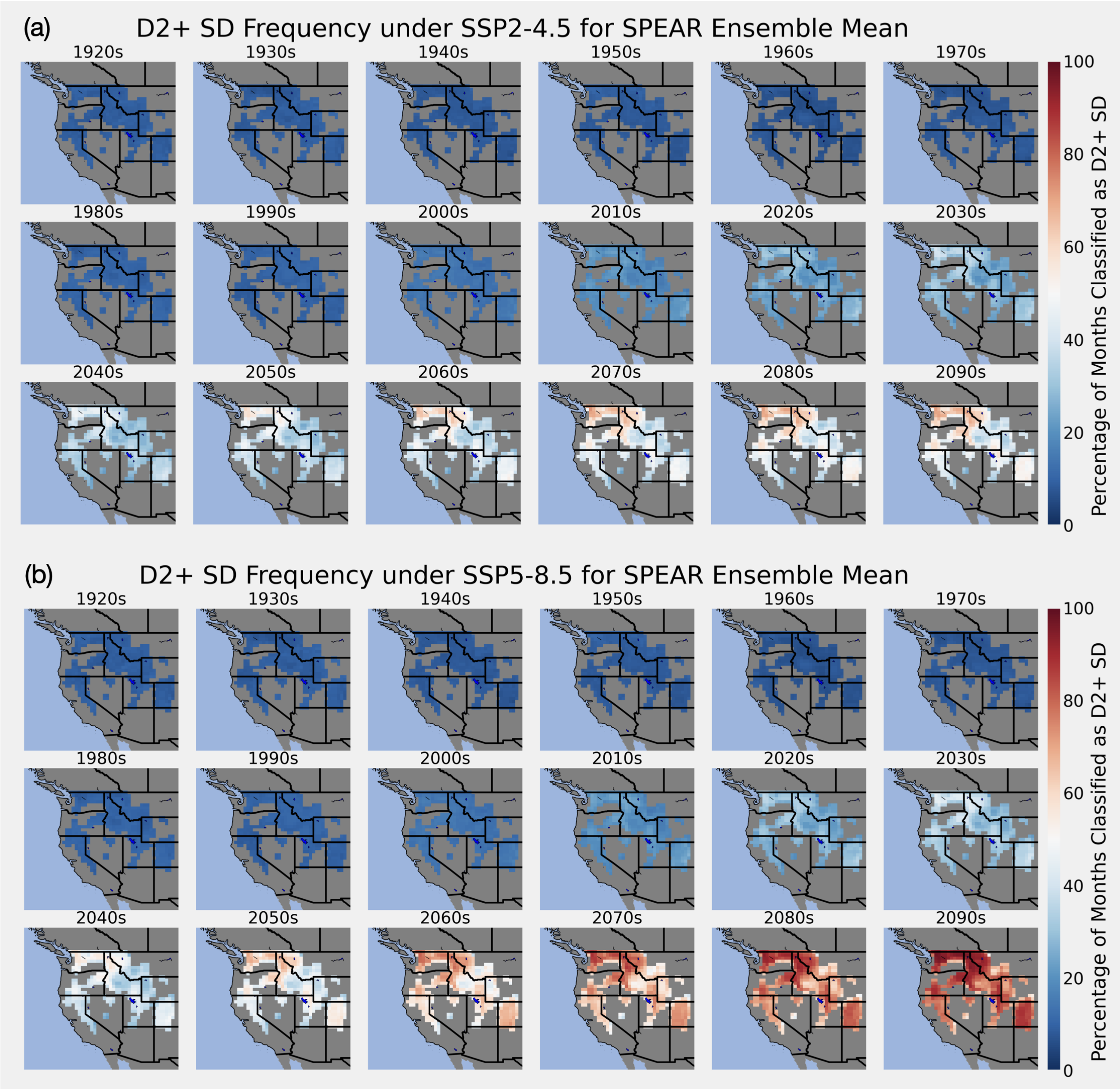

Figure S4. Panel plots for all 18 study decades between 1920 and 2100 for the SSP2-4.5 and SSP5-8.5 D2+ SD classification frequencies for the SPEAR ensemble mean. This figure emphasizes just how dramatic SPEAR projects the increase in D2+ SD occurrence to be, conditioned on the emissions scenarios, as the historical variability of the 20th century is barely distinguishable when placed on the same color scale as changes in the 21st century.

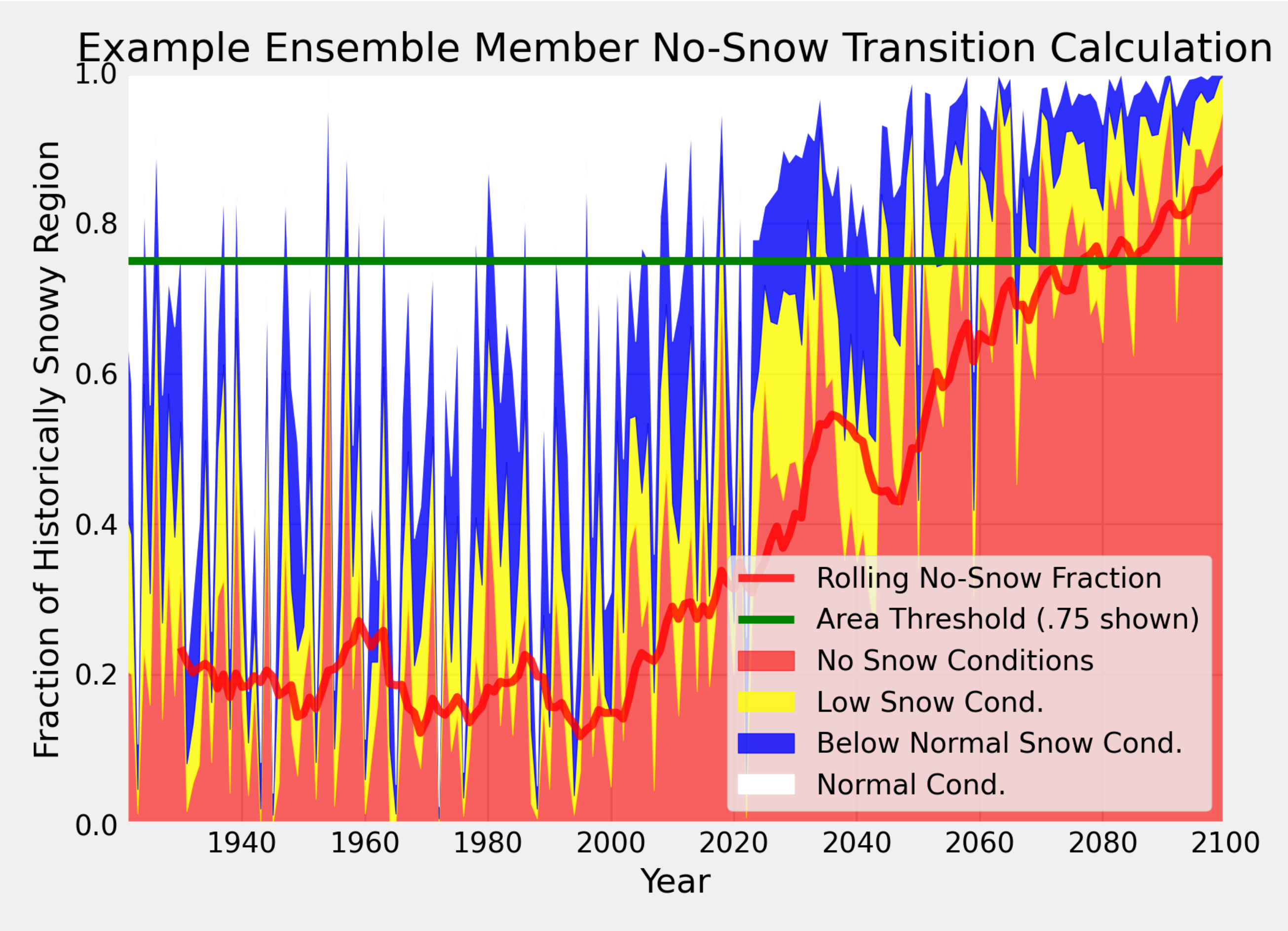

Figure S5. An illustration of how the no-snow transition is calculated as a function of the area threshold. This figure shows the fraction of the historically-snowy region experiencing no-snow (red), low-snow (yellow), below average-snow (blue), and near-normal or above-average snow (white) following the categories used by Siirila-Woodburn et al. (2021). The dark red curve represents a 10-year moving average of the yearly no-snow values (in red), while the green horizontal line indicates the chosen area threshold, in this case $\mathcal{A}$ = .75. For this particular region in one ensemble member, we see that the red curve crosses the green line for the last time in 2082. Thus, this ensemble member records a no-snow transition time of 2082 for the given threshold.