BAMS Article

Observing, Measuring, and Assessing the Consequences of Snow Drought

Alexander R. Gottlieb and Justin S. Mankin

ABSTRACT: Warmer and shorter winters from climate change will reduce snowpacks in most seasonally snow-covered regions of the world, with consequences for freshwater availability in spring and summer when people and ecosystems demand water most. Recent record-low snowpacks, such as those in the winters of 2013/14 and 2014/15 in the western United States, have led to a surge in research on "snow droughts," which are pointed to as harbingers of global warming that pose significant societal hazards. Yet, despite the importance of understanding snow droughts to best prepare for their attendant impacts, the concept remains amorphous, with no agreed-upon definition of what they are, how best to measure them, and how such snow droughts connect to warm-season impacts. These knowledge gaps limit our understanding of the risks posed by snow droughts in the present and future, and thus our preparedness for their differential impacts on freshwater resources. To address these issues, we compile a hemispheric ensemble of in situ, satellite, and reanalysis snowpack datasets. We use this ensemble to evaluate the scientific challenges and uncertainties arising from differences in defining and measuring snow droughts, and to identify opportunities to leverage this information to better understand the significance of snow droughts. We show that a clearer quantification of what constitutes a snow drought, including its uncertainties, improves our ability to anticipate costly and disruptive warm-season droughts, which is vital for informing risk management and adaptation to changing snow regimes.

KEYWORDS: Drought; Snowpack; Ensembles

https://doi.org/10.1175/BAMS-D-20-0243.1

Corresponding author: Alexander R. Gottlieb, alexander.r.gottlieb.gr@dartmouth.edu

In final form 10 November 2021

AFFILIATIONS: Gottlieb—Program in Ecology, Evolution, Environment and Society, and Department of Geography, Dartmouth College, Hanover, New Hampshire; **Mankin**—Department of Geography, and Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire, and Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

Snow is widely regarded as a vital component of the Earth system, with unique radiative and hydrologic properties that force modes of atmospheric circulation (Henderson et al. 2018; Zhang et al. 2019), shape the surface energy balance (Groisman et al. 1994), and serve as a natural reservoir that accumulates water during the winter and releases it during the warm season when human and ecosystem demands are highest (Barnett et al. 2005; Mankin et al. 2015). In particular, this role as a reservoir means snow can exert a strong influence on climate well beyond the winter, potentially shaping costly and disruptive warm-season climate extremes like droughts (Livneh and Badger 2020; Potopová et al. 2016), heat waves (Hall et al. 2008), or wildfires (Abatzoglou and Williams 2016; Westerling et al. 2006; Westerling 2016).

Given the perceived importance of snow to people and ecosystems, recent years of extremely low snowpack in some regions have raised concerns in the snow research and operations communities. The winters of 2013/14 and 2014/15, for example, saw widespread record-low spring snowpack measurements across the western United States that occurred in the context of dramatic long-term regional declines (Mote et al. 2018). The recent winter of 2020/21 was consistent with this trend, as many western U.S. basin snowpacks were below normal, reinforcing the exceptional drought that established itself in the summer of 2020. Similar trends toward smaller spring snowpacks have been observed across much of midlatitude Europe (Bach et al. 2018; Marty et al. 2017) and in the northeastern United States (Hodgkins and Dudley 2006). Perhaps unsurprisingly, these snowpack declines are partly attributable to anthropogenic warming (Berg and Hall 2017; Mote et al. 2018; Pierce et al. 2008). While internal climate variability generates uncertainty in the magnitude and direction of short-term snowpack trends, particularly in the high latitudes and cold continental interiors (Mankin and Diffenbaugh 2015; Siler et al. 2019), warming is expected to reduce snowpacks in most of the world's snow-covered regions (Bintanja and Andry 2017; Kapnick and Delworth 2013; Krasting et al. 2013; Pierce and Cayan 2013), making these currently unprecedented anomalies increasingly commonplace (Berg and Hall 2017; Diffenbaugh et al. 2013; Marshall et al. 2019).

Concern over these low snow conditions and the sense that they are a harbinger for a warmer future have inspired a spate of research on "snow droughts." This body of work wrestles with characterizing the historical frequency and intensity of low snowpacks (Huning and AghaKouchak 2020), their paleoclimatic context (Harley et al. 2020), their various meteorological causes (Harpold et al. 2017; Hatchett and McEvoy 2017), their anthropogenic fingerprint (Cooper et al. 2016; Huning and AghaKouchak 2018; Mote et al. 2016), and their anticipated frequency and intensity in the future under different warming scenarios (Marshall et al. 2019). Yet, while this research is largely motivated by concerns about the impacts of snow droughts in the present and in a warmer future, a number of hurdles have precluded rigorous assessment of the risks they pose to people and ecosystems in basins around the world. First, much of the geophysical analysis of changing snowpacks has been disconnected from their downstream consequences, as snow is implicitly assumed to be ubiquitously important wherever it contributes to runoff (Barnett et al. 2005). But the effects of snow droughts will be context dependent and shaped by differential factors like the role that snow plays in the regional water supply portfolio (Mankin et al. 2015; Qin et al. 2020). Only recently have analyses sought to concretely link the effects of global warming on snow's accumulation and melt to the sets of impacts that must be managed by people and ecosystems downstream (e.g., Berghuijs et al. 2014; Mankin et al. 2015; Qin et al. 2020; Viviroli et al. 2007). Second, there are enormous uncertainties in assessing snow droughts because of challenges in observing snowpack and a lack of agreement on snow drought metrics and the spatial and temporal scales over which to define them.

Here, we evaluate these barriers and highlight strategies that help our ability to assess the warm-season impacts of snow droughts. While the cold and shoulder seasons may also see impacts from snow droughts, including economic losses in communities that rely on snow-based recreation and midwinter reservoir management challenges (Barnett et al. 2005; Harpold et al. 2017), we focus on the potential warm-season impacts of snow droughts, using warm-season drought as a case study. Warm-season droughts rate among the most costly and disruptive climate extremes, causing economic damages of upward of $15 billion USD per year over the last decade in the United States (Smith 2020) and widespread ecosystem disruptions, including drought-induced forest mortality and elevated wildfire risk (Allen et al. 2015; Hartmann et al. 2018; Stephens et al. 2018).

To characterize observational and definitional uncertainties in snow drought, we compile a hemispheric ensemble of in situ, satellite remote sensing, and reanalysis snowpack datasets to which we apply common definitions from the snow drought literature. Using these data, we quantify the level of agreement of each dataset-definition combination at both the gridcell and hydrologic basin scales. We then consider how these uncertainties compound to shape our ability to predict the downstream consequences of snow drought, and identify empirical strategies for managing these uncertainties to glean more scientifically robust and operationally valuable insights about snow droughts. To illustrate this, we examine a case study of one potential consequence of snow droughts: an increased likelihood of warm-season droughts in regions like the western United States.

Our work emphasizes two points that are crucial to informing more effective adaptations to a less snowy future: first is the importance of fully characterizing uncertainties in estimates of snowpack at impacts-relevant scales and in the metrics and thresholds applied to those estimates to monitor snow droughts. Second is that observational uncertainty can be managed scientifically and made more meaningful to adaptation decision-making by concretely linking snow droughts to their warm-season impacts. We show that not only can we draw robust conclusions about the consequences of snow droughts despite the aforementioned challenges with observing and defining them, but we can also leverage these uncertainties to produce better forecasts of warm-season drought.

Defining snow drought

What is a snow drought? There is currently no agreed-upon community definition of snow drought. Instead there are a number of snowpack metrics and thresholds used, characterized at a wide range of spatial and temporal scales. While this definitional plurality is not intrinsically problematic, it does introduce uncertainty, as these analytical choices result in different quantifications of the frequency, intensity, and duration of snow droughts, and thus the sets of impacts we might attribute to them.

To illustrate this definitional uncertainty, we look at the evolution of the snowpack in the Columbia River basin during water year (WY, October–September) 2014/15, which was widely considered to be an historic snow drought in the Pacific Northwest (Harpold et al. 2017; Cooper et al. 2016; Harley et al. 2020) (Fig. 1a). We choose this event because the uncertainty across definitions of snow drought should be small, making it a useful benchmark for the possible level of agreement across definitions. As an indicator of snow

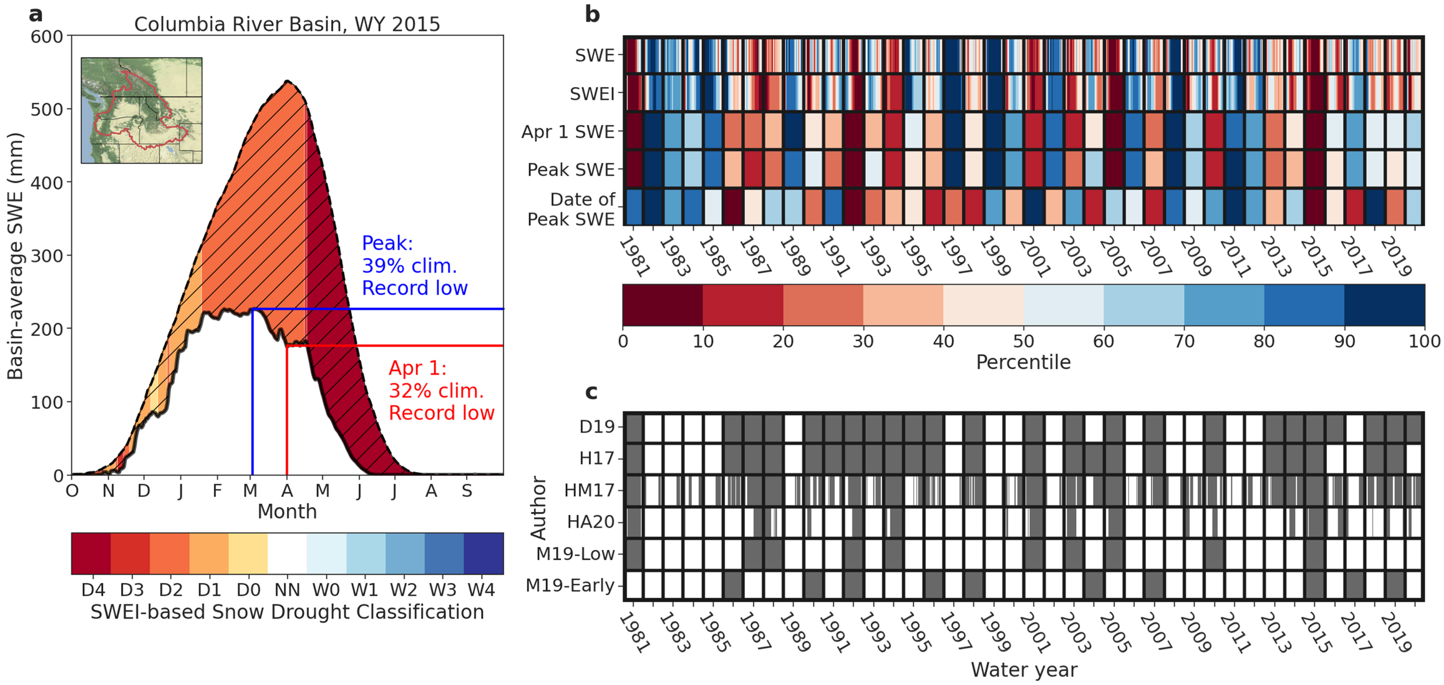

Fig. 1. (a) Common definitions of snow drought from the literature applied to SNOTEL data from WY 2014/15 in the Columbia River basin. For the area-weighted average across Columbia basin SNOTEL sites, the dotted line indicates daily climatological median snowpack while the solid line indicates the daily snowpack in WY 2014/15. This example illustrates different ways to assess that year's snow drought: 1) hatching highlights periods of SWE below the climatological median (Hatchett and McEvoy 2017); 2) color contours indicate snow droughts as diagnosed by Huning and AghaKouchak (2020) using the USDM D-scale; 3) the date and magnitude of peak SWE ("Peak," in blue); and 4) the magnitude of 1 April snowpack ("Apr 1," in red). (b) The level of agreement across common snowpack metrics across years, measured by the percentile value of the metric for the Columbia basin using SNOTEL data. (c) Identification of historical snow droughts (in gray) based on definitions from the literature [Dierauer et al. (2019): peak SWE < climatological mean; Harpold et al. (2017): 1 Apr SWE < climatological mean; Hatchett and McEvoy (2017): daily SWE < climatological median; Huning and AghaKouchak (2020): SWEI <–0.8; Marshall et al. (2019): Low – peak SWE < 25th percentile, Early – date of peak SWE < 25th percentile].

accumulation and melt in the Columbia that season, we examine daily data from the widely used network of Snowpack Telemetry (SNOTEL) sites, comparing WY 2014/15 to 40 years of the SNOTEL record in the basin spanning WY 1981–2020.

The data presented in Fig. 1 illustrate the uncertainty in snow droughts solely from how they are defined, even when using a single high-quality snowpack dataset, like in situ SNOTEL data. Before the deployment of automated snow pillows like SNOTEL in the late 1970s, SWE measurements came almost exclusively from manual snow surveys, which were traditionally performed at a monthly frequency, on or around the first of each month. Of these measurements, SWE on 1 April was considered to most closely approximate the maximum accumulation of snowpack, and thus the division between snow accumulation and melt seasons (Pagano et al. 2004). As such, 1 April SWE continues to be widely used in evaluations of snowpack and snow drought, as its lengthy observational record allows for the identification of long-term trends and variability (Mote et al. 2005, 2018). In 2015, the Columbia's 1 April SWE was an alarming 32% of normal. Despite the long precedent of evaluating winter snowpacks based on 1 April SWE, recent work has challenged its usefulness on the basis that snowpacks in different regions can peak on average over a month before (e.g., the Pacific Northwest) or after (e.g., cold continental regions such as the Northern Rockies) 1 April. Across western North America as a whole, roughly a quarter of accumulated snowpack has already melted out by 1 April (Musselman et al. 2021). As such, others have advocated for the use of peak SWE and/or its timing as a potentially more hydrologically relevant quantity, as it represents the maximum amount of water available for runoff generation and soil moisture recharge at the start of the melt season (Marshall et al. 2019). During the 2014/15 Columbia snow drought, peak SWE was the lowest in the SNOTEL record at 39% of the climatology, occurring around 1 March (Fig. 1a).

Point-in-time measurements like 1 April or peak SWE are fundamentally retrospective characterizations of the snow season. Because of this, such measures have limited ability to provide early warnings of snow droughts, which reduces the time to prepare for their impacts. Hatchett and McEvoy (2017) and Huning and AghaKouchak (2020) consider snow droughts as subseasonal phenomena that can begin, intensify, and abate all within the course of a snow season. We show two examples of subseasonal snowpack measures from the respective authors: daily SWE relative to the climatological median (Fig. 1a, hatching) and the SWE index (SWEI), the latter of which is defined based on a 3-month rolling sum of daily SWE (Fig. 1a, shading). It is clear that these continuous indices are more conducive to active monitoring of snow droughts. Such indices are also useful for finer-scale analyses of the different impacts snow droughts might generate based on when within the snow season they occur (Hatchett and McEvoy 2017). Both continuous indices also emphasize the deepening of the Columbia River basin into a more intense snow drought as the season progresses, and show this considerably earlier than the point-in-time diagnostics of peak or 1 April SWE (Fig. 1a).

Comparing these metrics of snow states across time for the Columbia River basin emphasizes their similarities and differences. While the measures generally agree on the qualitative direction of snow anomalies in a given WY (columns in Fig. 1b), they give different answers about the timing and intensity of snow drought events. Hatchett and McEvoy (2017), for instance, would identify the onset of the WY 2014/15 snow drought when SWE first dipped below the climatological median, effectively from the very start of the WY in October. In that same year, Huning and AghaKouchak's (2020) SWEI would not identify a continuous snow drought event (D1 or worse) until January. Furthermore, we can get very different impressions of the severity of the 2014/15 event depending on whether we look at the historically low peak and 1 April SWE or the SWEI, which would indicate only a severe drought (D2) at their respective points in time, only progressing to an exceptional drought (D4) roughly a month later (Fig. 1a).

In addition to metric choice, the snow drought thresholds applied to those metrics can limit the consistency between authors on the diagnosis of any one event (Fig. 1c). The 2014/15 snow season is but one of two years in the record (the other being WY 1992/93) where all approaches we consider here agree on a snow drought classification using the same dataset. We emphasize that identifying when and where different definitions of snow drought agree or disagree is not a pedantic exercise: these analytical choices inform our understanding of the spatiotemporal patterns of snow droughts and their characteristics, and present an additional hurdle to understanding their historical and anticipated warm-season impacts. This is because any given analysis will be starting from a fundamentally different set of events and could therefore lead to fundamentally different conclusions.

Finally, we note that some have recently advocated for defining snow droughts based not just on snowpack, but on their underlying causes. For example, a typology that differentiates between precipitation-limited, or "dry," snow droughts, and their "warm" counterparts where high temperatures prevent near- to above-normal precipitation from accumulating as snowpack, has been gaining traction (Harpold et al. 2017; Hatchett and McEvoy 2017; Marshall et al. 2019). The 2014/15 Pacific Northwest snow drought examined here would be classified as warm, with sufficient precipitation, but extremely high winter temperatures (Cooper et al. 2016; Harpold et al. 2017). While we do not explicitly employ such a typology here or assess the differential impacts of snow droughts based on their drivers, this framework represents a step toward defining snow droughts based in part on their anticipated consequences and the challenges they might present. For instance, reservoir managers are tasked with the balancing act of maximizing the capture of inflows while maintaining enough excess capacity to accommodate future heavy flow events (e.g., during spring snowmelt pulses). While a warm snow drought may require difficult management decisions about whether to retain or release large midwinter flows from runoff from rainfall that might normally have been stored in snowpack and/or early melt events, a dry snow drought may not have the same inflow management implications, but may exert greater warm-season stress on streamflow and reservoir levels (Harpold et al. 2017). We return to this idea of defining snow droughts based in part on their likely differential consequences in our concluding remarks.

The observational basis for snow drought

While in situ point measurements of SWE such as the SNOTEL data presented in Fig. 1 or manual snow course surveys are valuable for snow science and water supply forecasting in western North America, the ability to draw robust conclusions about snow drought and its impacts from them is limited by a few factors. First, an observational record of in situ measurements long enough to track long-term trends and variability in SWE is available only in a limited number of river basins in western North America, limiting conclusions about spatiotemporal patterns in snowpacks and their downstream impacts to a small set of hydroclimatic (and water resource management) contexts. While snow course data provide the long-term continuous observations of western U.S. SWE required to assess recent snow seasons relative to their climatology (e.g., Mote et al. 2018), their low sampling frequency may miss hydrologically important quantities, such as peak SWE and its timing, as well as subseasonal accumulation and melt dynamics (Musselman et al. 2021). Such sampling gaps also prevent the use of snow drought definitions (like SWEI) that allow for subseasonal identification of snow droughts and provide more lead time to manage their impacts. Even areas fortunate enough to have a high density of daily in situ data suffer from issues of representativeness that limit their generalizability and the calculation of hydrologically important quantities such as basin-scale SWE (Meromy et al. 2013; Molotch and Bales 2006; Mortimer et al. 2020). These issues have implications for providing accurate snow drought impacts analyses, as many of the consequences of snow drought occur downstream from where snow accumulates (Barnett et al. 2005; Viviroli et al. 2007) and is measured.

Accordingly, snow droughts should be considered at spatial scales commensurable with their anticipated impacts, such as hydrologic basin or subbasin scales, which requires gridded data. Satellite remote sensing provides spatially continuous estimates of SWE that can be aggregated to these scales, but the coarse resolution of these data products misses considerable variation in SWE due to terrain and land cover heterogeneities (Mortimer et al. 2020). Furthermore, estimates based on passive microwave retrievals, which provide the longest-running record of spatially continuous daily SWE over a hemispheric or global domain, suffer from issues of interference from vegetation and a saturation effect that makes it difficult to accurately quantify deep snowpacks (Dietz et al. 2012). Finally, reanalyses that statistically or dynamically estimate SWE allow for spatiotemporally consistent and high-resolution estimates of snowpack, but these products are highly sensitive to the forcing data and snow physics schemes they employ (Kim et al. 2021; Mudryk et al. 2015). These issues notwithstanding, both satellite remote sensing observations and reanalysis products are difficult to validate against in situ data. For example, it is not clear how best to compare gridded products whose spatial scales represent an average over tens to tens of thousands of square kilometers of complex terrain against a small sample of point measurements, should any be available (Mortimer et al. 2020). As such, there is little ground truth knowledge of the size and state of snowpacks beyond the point scale for vast snow-covered regions.

Remote sensing products and reanalyses that necessarily constitute the observational basis for basin, regional, hemispheric, or global snow drought analysis are inherently uncertain. We characterize this uncertainty by compiling an ensemble of 15 gridded SWE data products that cover the Northern Hemisphere (Table 1), each regridded to a common

Table 1. Summary of SWE data products used in the observational ensemble.

Dataset	Method	Forcing data, if applicable	Spatial resolution	Spatial coverage	Temporal resolution	Temporal coverage	Reference
ESA GlobSnow v3.0	Satellite passive microwave + in situ	—	25 km	NH	Daily	1979–2018	Luojus et al. (2020); Pulliainen et al. (2020),
NASA AMSR-E/AMSR2	Satellite passive microwave + in situ	—	25 km	Global	Daily	2002–present	Tedesco and Jeyaratnam (2019)
Canadian Meteorological Centre (CMC)	Global Environmental Multiscale (GEM) Model + in situ	GEM	35 km	NH	Daily	1998–present	Brown and Brasnett (2010)
ERA5-Land	H-TESSEL LSM	ERA5	0.1° × 0.1°	Global	Hourly	1981–present	Sabater (2019)
NASA GLDAS 2.0	Catchment, Noah, and VIC LSMs	Princeton	0.25° × 0.25°	Global	3-hourly	2000–present	Rodell et al. (2004)
NASA MERRA-2	Catchment LSM	MERRA-2	0.5° × 0.625°	Global	3-hourly	1980–present	Gelaro et al. (2017)
JRA-55	Simple Biosphere Model (SiB)	JRA-55	55 km	Global	6-hourly	1958–present	Kobayashi et al. (2015)
NASA DayMet V4	Temperature-driven accumulation/melt model	DayMet	1 km	North America	Daily	1980–present	Thornton et al. (2020)
NASA NLDAS	Mosaic, Noah, and VIC LSMs	NARR, CPC	0.125° × 0.125°	CONUS	Hourly	1979–present	Xia et al. (2012)
NOHRSC SNODAS	Snow mass and energy balance model + satellite, airborne and in situ obs.	RUC2 NWP	1 km	CONUS	Hourly	2003–present	National Operational Hydrologic Remote Sensing Center (2004)
University of Arizona (UAZ)	Statistically interpolated in situ	PRISM	4 km	CONUS	Daily	1981–present	Broxton et al. (2019); Zeng et al. (2018)

0.5° × 0.5° resolution. We use this observational ensemble to calculate a number of snow drought metrics and evaluate the level of ensemble agreement on those measures. For instance, Fig. 2 shows the observational uncertainty in one diagnostic for evaluating snow drought, the magnitude of peak SWE. Across most of the hemisphere (save a few small regions of the Canadian Arctic and Siberia), the average ensemble spread of peak SWE exceeds 100% of the ensemble mean (Fig. 2a, red colors). This level of disagreement suggests that gridded observations do not provide a reasonable degree of certainty about the magnitude of snowpack in a given location and year. Such a finding is consistent with the expectation of significant biases in individual data products (Kim et al. 2021; Mudryk et al. 2015; Mortimer et al. 2020), and emphasizes the potential sensitivity of research claims about SWE magnitudes to data choices.

Because of the observational uncertainty in the true size of snowpacks, most snow drought analyses instead focus on relative measures, like snowpack anomalies or their ranks within a single dataset (Cooper et al. 2016; Huning and AghaKouchak 2020; Marshall et al. 2019; Mote et al. 2016). Here the implicit assumption is that even if any one data product cannot reliably estimate the true snowpack magnitude, there should be broad agreement on how snowpacks vary through time. Yet the dark red colors in Fig. 2b show that this is not the case. The average correlation of peak SWE anomalies among any two datasets we consider is low (<0.5) for most grid cells in the Northern Hemisphere. This finding suggests that in addition to not knowing the true local magnitude of snowpack, we cannot confidently determine whether snowpack in that year is above or below normal. As such, using gridded snowpack anomalies or their rankings as a means to assess the occurrence or severity of snow droughts may also suffer from a sensitivity to data choices.

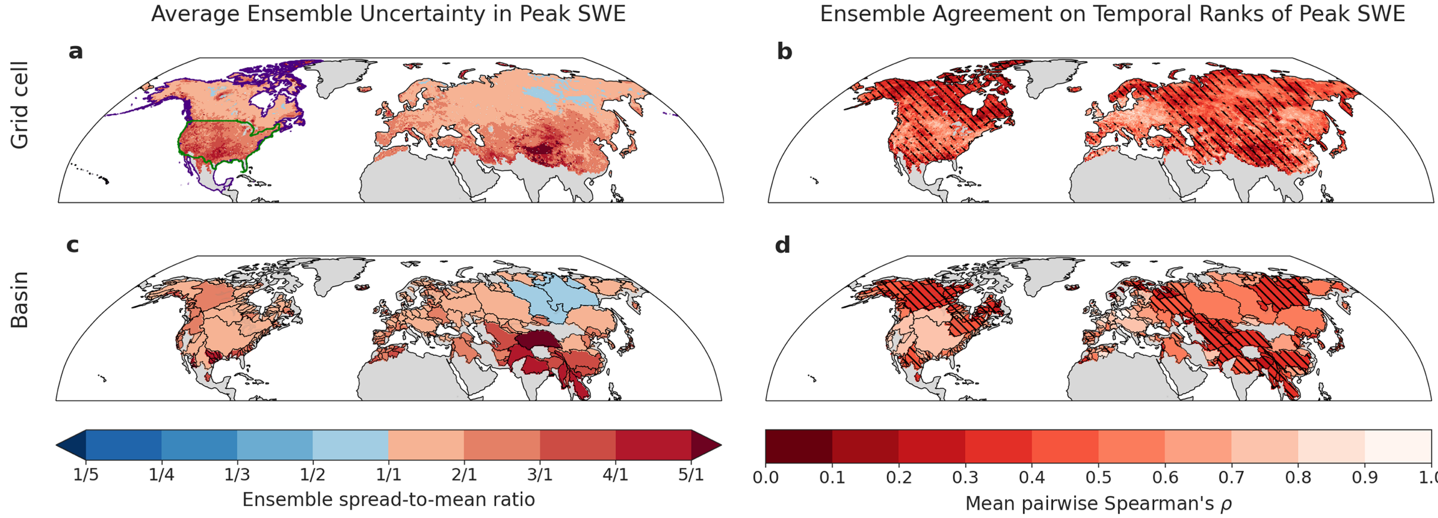

Fig. 2. (left) Average ratio of the within-ensemble range in annual peak SWE to the ensemble mean at the (a) gridcell and (c) basin scales. (right) Average pairwise Spearman's rank correlations of peak SWE time series across the ensemble at the (b) gridcell and (d) basin scales. Darker reds indicate weaker agreement between any two datasets in the ensemble; hatching indicates regions where fewer than half of all pairwise correlations are statistically significant ($p < 0.05$). Regional outlines in (a) indicate the number of data products used in the ensemble (green = 15, purple = 10, no outline = 9).

The issue of uncertainty in snowpack magnitude and variability persists even if we consider them at larger, more hydrologically meaningful scales, such as the river basin. The overwhelming majority of basins show average ensemble uncertainty in peak SWE of greater than 100% (Fig. 2c). And while the observations track year-to-year variability in basin-scale snowpack more consistently in a number of basins in the United States and central Europe than at the gridcell level (cf. Figs. 2b,d), many of the major snow-dependent basins in the high latitudes and high-mountain Asia have little agreement on the direction of snowpack anomalies across time (Fig. 2d). The lack of in situ data for many snow-reliant regions, coupled with the need to make snow drought assessments at the basin, regional, hemispheric, or global scales, forces the use of gridded products. But our observational ensemble of hemispheric snowpack highlights new challenges this solution brings: when comparing gridded datasets, it is difficult to make robust claims about the direction or magnitude of true snowpack anomalies across space and time. Such uncertainty makes it potentially misleading to rely on a single dataset to make claims about snow droughts or their impacts. Together, these challenges raise questions about how best to characterize and analyze snow droughts. We consider one way to manage this observational uncertainty: using an ensemble to connect snow droughts to their impacts.

Connecting snow drought to warm-season drought

Despite the uncertainties in what constitutes a snow drought (Fig. 1) and in snowpack measurements beyond the in situ scale (Fig. 2), there remains an imperative to understand snow droughts and predict their warm-season impacts. Concerns about the implications of snow drought for warm-season drought are a primary motivation in the snow drought literature and are underpinned by the demonstrated effects of snowpack variability on the timing and magnitude of streamflow (Stewart et al. 2005; Barnett et al. 2005; Berghuijs et al. 2014), soil moisture (Hall et al. 2008; Harpold and Molotch 2015), and ecosystem water use (Molotch et al. 2009). Accordingly, we ask, "Does information about snow drought, however measured and defined, actually improve forecasts of droughts in the warm season?"

To answer this question, we quantify how much forecasts of warm-season drought improve when we include information about snow drought occurrence in the preceding snow season. We assess this improved skill relative to a simple (null) model that uses only temperature and precipitation from the first half of the WY as predictors. The dependent variable is a negative change in drought classification during the warm season (i.e., a drought onset or intensification), as defined by the U.S. Drought Monitor (USDM). The USDM is a synthetic drought product that combines geophysical measurements, expert judgment, and local networks of observers of the impacts of drought to identify droughts of five levels of severity, from abnormally dry (D0) to exceptional drought (D4). The USDM's multivariate nature is particularly useful for considering the impacts of snow drought, as snow may shape warm-season conditions through different pathways in different regions. For instance, the timing and magnitude of runoff from snowmelt may be the primary concern in mountainous regions with highly seasonal precipitation (Viviroli et al. 2007), while its contribution to spring soil moisture recharge and subsequent land–atmosphere interactions may be a more important driver of drought in the Great Plains (Hall et al. 2008). By taking into account the regional context of drought, the USDM classifications are a more robust indicator of truly impactful drought events than single variable measures (e.g., runoff or soil moisture) and are less sensitive to observational uncertainty in those quantities.

Crucially, information about snow drought dramatically improves the prediction of warm-season droughts in western U.S. basins, even in the face of the large definitional and data uncertainties we present in the preceding sections (Fig. 3). Moreover, we show that snow drought information from the entire ensemble of dataset–definition combinations is more skillful in predicting warm-season drought than that from any dataset–definition combination alone, including that from in situ SNOTEL. The colored markers in Fig. 3a show the improvement in warm-season drought forecast skill for example basins that comes from including particular dataset–definition combinations (Table A1) from our observational snowpack ensemble. The black markers show the forecast skill improvement that comes from the including SNOTEL data in our empirical warm-season drought prediction. Forecast skill for each dataset–definition combination is assessed relative to a null model that uses only temperature and precipitation from the first half of the WY (October–March) as predictors (Fig. 3b).

We emphasize a few key findings from the analysis presented in Fig. 3: first is that the concept of snow drought, however measured and defined, is generally useful for forecasting warm-season drought, save for those basins in which the forecast skill of the null model is already quite high (e.g., the Rio Grande, Rogue, and Skagit). Our finding is consistent with previous results showing the importance of snowpack as a source of skill for forecasting warm-season streamflow (Koster et al. 2010; Livneh and Badger 2020). Second, at the same time, our understanding of the relationship between snow droughts and their warm season counterparts is sensitive to how snow droughts are measured and defined. This sensitivity can be seen in the wide range of forecast skill that is gained by including any one dataset–definition combination into the null model. Some dataset–definition combinations offer little improvement in skill while others more than double it. Finally, there does not appear to be any particular snowpack dataset, definition, or combination thereof that is consistently a top performer across basins, further emphasizing the limitations of relying on a single snowpack dataset and means of identifying snow droughts.

Yet, this disagreement does not mean that we cannot draw robust conclusions about the relationship between snow drought and warm-season drought in these basins. Rather, we can leverage the observational and definitional uncertainty across the ensemble to produce forecasts that exceed the skill of any one member alone. The blue lines in Fig. 3 represent estimates of the percent change in warm-season drought forecast skill from including the ensemble-mean snow drought classification (i.e., the fraction of dataset–definition combinations that identify a snow drought in a given year). We weight each member by its forecast skill, and present the median estimate as a contrasting white line (see the appendix). For all western U.S. basins, save for the Skagit, Rogue, and one instance in the San Joaquin, use of the ensemble mean increases warm-season drought forecast skill more than any individual

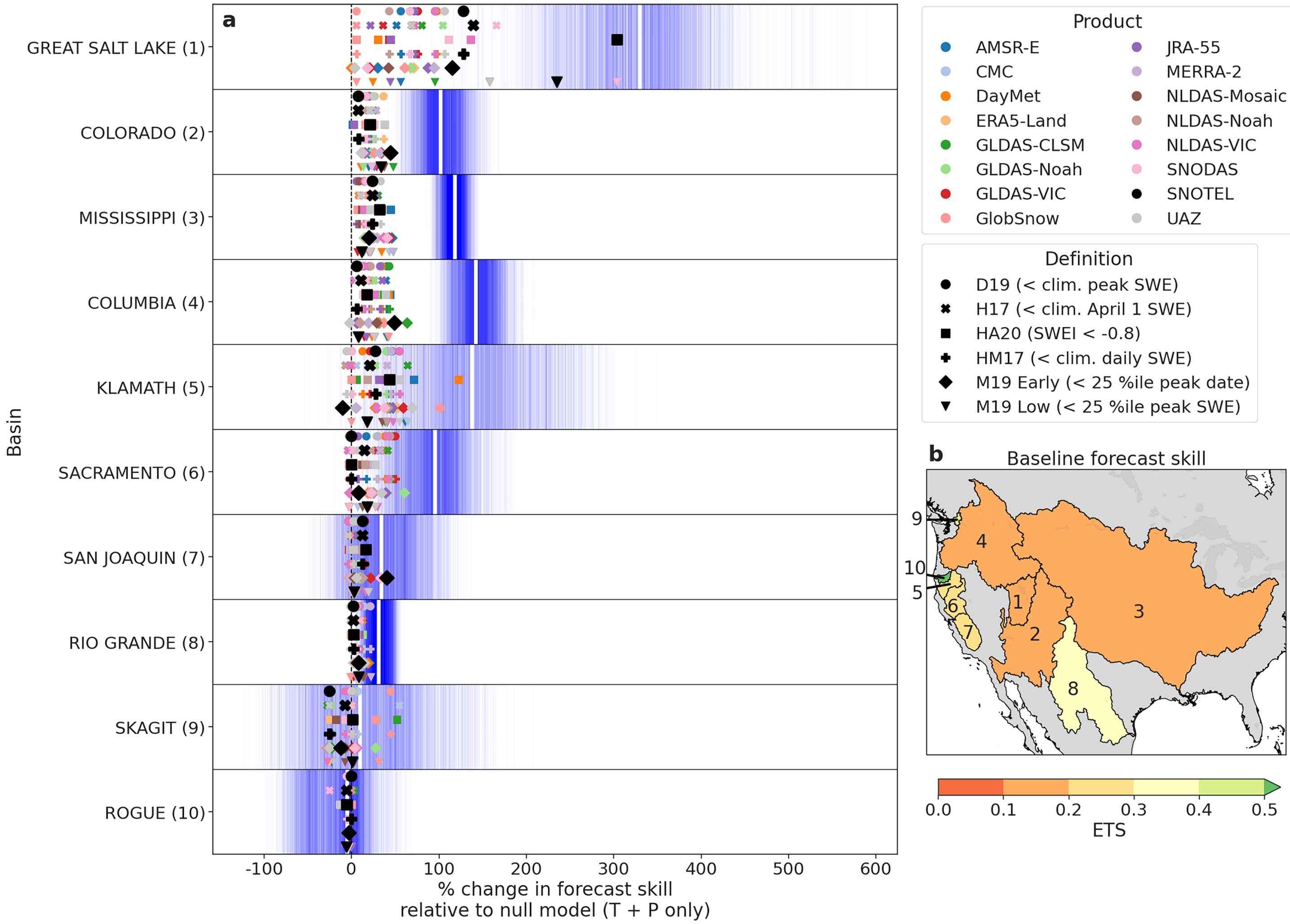

Fig. 3. (a) Basinwide percentage change in forecast skill (based on the ETS) of warm-season drought from inclusion of individual dataset–definition combinations (markers) relative to a null model based on seasonal temperature and precipitation alone (see the appendix). The colored markers represent the values from individual ensemble members, while the black markers represent the in situ SNOTEL data. The blue lines represent repeated *K*-fold cross-validation estimates of the change in forecast skill with the inclusion of the ensemble-mean snow drought classification (i.e., the fraction of dataset–definition combinations that identify a snow drought in a given WY), weighted by each individual dataset–definition combination's ETS. The median change in forecast skill with the ensemble is indicated by a white line. Basins in (a) are ordered based on the skill of the null model, with the top (Great Salt Lake) having the lowest baseline ETS. All changes in ETS are expressed relative to the map presented in (b), which shows the forecast skill of the null model (which uses only average temperature and cumulative precipitation from the first half of the WY (October–March) to forecast warm-season drought onset/intensification, as determined by the USDM).

ensemble member. For major basins, like the Colorado, Mississippi, and Columbia, the forecast skill improves by a factor of 2 or more relative to the null model in Fig. 3b. Moreover, the greatest relative increases in forecast skill are seen in basins with lower predictability from temperature and precipitation alone (warmer colors in Fig. 3b corresponding to top rows of Fig. 3a). We emphasize that these results do not indicate the strength (or lack thereof) of the underlying biogeophysical connection between snow drought and warm-season drought; rather, they indicate where snow droughts can provide the most additional information about the likelihood of this type of warm-season impact.

Our basin-by-basin, dataset-by-dataset, definition-by-definition approach to analyzing the relationship between snow droughts and warm-season droughts highlights at least three points relevant to snow drought research: first, it provides an objective basis to evaluate the myriad datasets and definitions of snowpack by linking snow drought to an impact of interest. While we do not undertake the exercise of identifying the best-performing ensemble members, an ensemble like this provides analysts a basis for identifying which datasets and definitions best identify particular snow drought consequences. Beyond the in situ scale, we may not be able to evaluate which observations are closest to truth, but we can determine which snowpack datasets and snow drought definitions provide the most information about the impacts of interest. This practice of mapping a large observational ensemble onto impacts and using the observed relationship to adjudicate between datasets and definitions may be particularly valuable in regions of the world where there is weaker agreement in the snowpack observations (e.g., high-mountain Asia in Fig. 2) and/or a lack of in situ data. Second, there is value in identifying the range and source of uncertainties in snow droughts and their impacts, as it provides potential paths to constrain those uncertainties and inform better decisions. For instance, the variability in forecast skill is generally greater across datasets than definitions (Fig. 3a), suggesting that constraining our snowpack observations might result in greater confidence in snow drought–based forecasts of warm-season drought than the identification of better snowpack metrics and thresholds of those metrics. Finally, while observational and definitional uncertainties might lead to different conclusions about the impacts of snow droughts, such uncertainties can be leveraged together to perform analyses that are both more operationally valuable and scientifically robust than could be achieved with any one particular data or definition choice alone.

Moving forward

We have illustrated the scope of definitional and data uncertainties related to snow drought and how those uncertainties can influence scientific claims about their impacts. Despite these challenges, we argue that there is value in fully quantifying and leveraging these uncertainties: for example, our ensemble of dataset–definition combinations of snow drought is a more skillful predictor of warm-season drought than even in situ data. We conclude by discussing the operational utility of this approach and highlighting a few promising avenues of research that will allow us to deepen our understanding of snow droughts and their consequences.

Initial hydrologic conditions (IHCs) such as snowpack and soil moisture have long been known to be an important source of skill in seasonal streamflow forecasting (Koster et al. 2010; Livneh and Badger 2020), and snowpack measurements in particular have underpinned water supply forecasts in the western United States for over a century (Pagano et al. 2004). Our findings suggest that considering all of the attendant uncertainties of IHCs may offer even more predictive benefit. For instance, the monthly and seasonal drought outlooks generated by the National Weather Service's Climate Prediction Center (www.cpc.ncep.noaa.gov/products/Drought/) use expert judgment guided by empirical and process-based ensemble forecasts of temperature and precipitation and current IHCs to determine probabilities of drought development, persistence, and abatement, using the USDM maps as initial conditions for each assessment. Our results suggest that there may be value in more quantitatively leveraging the uncertainties in IHCs such as snowpack when issuing warm-season drought forecasts in the spring. This could take the form of either direct integration into the dynamical and statistical forecast models already in operation that inform forecasts, or as a standalone statistical model, such as the one presented here, whose predictions can add an additional data source to guide expert judgment.

There are also a number of promising avenues of scientific research that can advance our understanding of snow droughts and the risks they present. First, we point to the well-vetted set of best practices on how to analytically manage ensemble uncertainties, such as those that have arisen in the climate modeling community, like the Coupled Model Intercomparison Projects (CMIP). The observational ensemble of snowpack we present here, like the CMIP, is one of opportunity, not design. That reality has implications about how we assess performance and skill and the statistical measures that are most appropriate to use on it. For example, while we can benchmark the various gridded data products against one another and against in situ point observations (Kim et al. 2021; Mortimer et al. 2020; Mudryk et al. 2015), we cannot distinguish which best capture the accumulation and melt of snow that actually occurred in the real world at the most scientifically or operationally relevant scales. Additionally, while we have shown that the mean across our observational and definitional ensemble outperforms most if not all constituent members, as is also generally true of the CMIP ensembles (Annan and Hargreaves 2011), the ensemble mean itself is likely to contain persistent biases (Annan and Hargreaves 2017). Thus, if neither the ensemble members nor the ensemble mean is a more robust estimate of the true snowpack, it may be more prudent to consider each member as statistically indistinguishable from truth (Annan and Hargreaves 2011), or to weight them based on the strength of their relationship to some other quantity of interest, as we have done here with warm-season drought in Fig. 3. We do not provide an exhaustive treatment of the practices from multimodel climate analysis that are transferable to an ensemble of snowpack observations, but suggest that insights from this community may inform how we approach the snow data for particular applications.

Second, our approach to assessing the consequences of snow drought is top-down in nature: prescriptive definitions are applied to datasets, and the resulting quantities are used to predict the impact of interest. We acknowledge, however, that there is value in bottom-up approaches that begin with impacts and identify the antecedent snowpack conditions that select for those impacts. Critically, such approaches can generate emergent definitions of snow drought based on specific impacts in specific locations and identify the most useful set of observations, metrics, and thresholds for assessing those impacts. Bottom-up approaches could also complement the dry versus warm snow drought typologies advocated by some authors (Harpold et al. 2017), as it would allow us to evaluate how the drivers of snow droughts [to include modes of variability, e.g., Cook et al. (2018) and McCabe and Dettinger (2002)] select for particular impacts in a wide variety of snow regimes. Furthermore, this type of analysis is easily extended to consider multiyear or chronic snow droughts and how they may compound to have unique, nonlinear effects on water resources (Marshall et al. 2019). While bottom-up analyses are contingent upon having reliable on-the-ground impacts data (which may not be available everywhere snow droughts are truly impactful) such approaches may be the clearest means of identifying the most concerning snow conditions.

Finally, while forecast analyses such as the one we present in Fig. 3 offer operationally relevant information to stakeholders, considerable work remains to understand the biogeophysical mechanisms through which snow influences warm-season climate. Existing work has associated regional-scale antecedent snow conditions with wildfire (Westerling 2016; Westerling et al. 2006) and here we have demonstrated a basin-scale association with warm-season droughts. However, as of yet, there has been limited assessment of the physical processes that link them. This is crucial, because understanding the full implications of a warmer future with more snow droughts requires understanding how their occurrence does or does not select for particular impacts. We have ample theory and findings from adjacent literatures to suggest mechanisms through which snow droughts could influence warm-season climate—for instance, through altering the timing and magnitude of streamflow (Hatchett and McEvoy 2017), shaping vegetation phenology, productivity, and water use (Cayan et al. 2001; Pulliainen et al. 2017; Trujillo et al. 2012), and promoting warming and drying land–atmosphere feedbacks through their influence on surface water and energy fluxes (Groisman et al. 1994; Kolstad 2017; Meira Neto et al. 2020; Molotch et al. 2009). But the entire network of associations between snow, the atmosphere, soils, vegetation, and runoff that might allow snow droughts to influence the warm season remains underexplored. Diagnosing these pathways for snow-dependent basins in the present climate is vital, as many of these potential linkages, such as the strength of land–atmosphere coupling

(Seneviratne et al. 2006) or the effects of plants on the surface water balance (Mankin et al. 2019; Ukkola et al. 2016) are also expected to evolve in a nonstationary climate. Thus, detailed process tracing utilizing both observations and idealized experiments such as those in the Land Surface, Snow, and Soil Moisture Model Intercomparison Project (LS3MIP) contribution to CMIP6 (van den Hurk et al. 2016) is necessary to identify not just where snow droughts have downstream consequences, but how.

Even under moderate future warming scenarios, cold-season precipitation will be increasingly partitioned toward rain over snow, snow accumulation will decrease, and the snow season will contract in the overwhelming majority of the world's seasonally snow-covered regions (Ashley et al. 2020; Mankin and Diffenbaugh 2015; Pierce and Cayan 2013). Accordingly, snow droughts, however defined, are bound to become increasingly commonplace. But their consequences will not be felt equally. Such snow droughts will have massive water availability implications in some basins, but not in others. Our ability to identify differential vulnerabilities and manage them is contingent upon our understanding of where, when, and how snow droughts select for impacts in human and natural systems. Such insights can only come through analyses that consider all sources of uncertainty, including ones about data and definition choices, as we have illustrated here. While the challenges from observing, characterizing, and predicting snow droughts are considerable, the practice of evaluating how our scientific choices influence our claims about them is a first step to inform more efficient and effective adaptations to a less snowy future.

Acknowledgments. We thank Dartmouth's Research Computing and Discovery Cluster for computational support, and each of the SWE dataset authors for their work providing snowpack data. The authors also thank NOAA's Yun Fan for helpful discussions about the NOAA Drought Outlooks forecast methodology. A.R.G and J.S.M. are supported by DOE Grant DE-SC0022302. J.S.M. is also supported by the Burke Research Initiation Award and NOAA MAPP Grant NA20OAR4310414.

Data availability statement. All data and code used in the analysis will be made publicly available at https://github.com/alex-gottlieb/snow_drought. All snowpack data used in the creation of the observational ensemble are openly available at locations cited in the reference section and Table 1, and instructions for downloading the raw data will be provided in the GitHub repository.

Appendix: Data and methods

We compile an ensemble of 15 satellite remote sensing and reanalysis datasets of snow water equivalent (SWE): 9 for the entire Northern Hemisphere, an additional dataset for all of North America, and 5 more for the continental United States only (Table 1). All SWE values are rescaled to the common units of mm (equivalently, kg m^{-2}). All datasets are regridded to a common 0.5° × 0.5° resolution using bilinear interpolation and datasets with a sub-daily native temporal resolution (ERA5-Land, MERRA-2, JRA-55, SNODAS, all 3 GLDAS products, and all 3 NLDAS products) are resampled to a daily resolution by taking the mean of all values within a given day. The Canadian Meteorological Centre SWE values are calculated

Table A1. Snow drought definitions used in the analysis.

Authors	Metric	Snow drought threshold
Harpold et al. (2017)	1 Apr SWE	Below climatological mean
Dierauer et al. (2019)	Peak SWE	Below climatological mean
Marshall et al. (2019)	Peak SWE	Below historical 25th percentile
Marshall et al. (2019)	Date of peak SWE	Before historical 25th percentile
Hatchett and McEvoy (2017)	Daily SWE	Below climatological median
Huning and AghaKouchak (2020)	Daily SWEI	<–0.8 (D1)

from the Daily Snow Depth analysis using the same lookup table of snow densities used in generating their monthly SWE product (Brown and Brasnett 2010). All analysis is performed on the shared time period among datasets, water years (WY, the 12-month period from the previous October to September) 2004–18, and all anomalies are calculated relative to this common period.

Given the strong imperative to consider snow droughts at scales relevant to hydrologic impacts, we aggregate SWE to the basin scale using basin extents from the Global Runoff Data Centre's Major River Basins of the World database (GRDC 2020). Before aggregating, gridcell-level average SWE is converted to total snow mass (in kg) by multiplying by the area of the grid cell (in m^2). Only basins where at least half of all product years register nonzero snow mass are used in the analysis.

From this ensemble of SWE observations, we calculate a number of snow drought metrics for each product at both the gridded and basin scales: peak SWE, or the maximum daily value of SWE within each WY; the ordinal date of peak SWE relative to the start of the WY; 1 April SWE; cumulative SWE deficit or surplus relative to median climatology; and a nonparametric SWE index (SWEI) from Huning and AghaKouchak (2020). The latter is calculated by taking a 90-day rolling sum of SWE and determining the probabilities of each value using the Gringorten plotting position (Gringorten 1963), given by

$$p=\frac{i-0.44}{N+0.12},$$

where i is the rank of the value for each day of the water year across all N years of data in the analysis. The standardized SWEI is calculated by transforming the empirical probability p using the inverse standard normal distribution:

$$\mathrm{SWEI}=\Phi^{-1}\left(p\right).$$

We characterize the observational uncertainty u in the magnitude of snowpack in Fig. 2 by calculating the difference between the ensemble maximum and minimum values of peak SWE, normalizing by the ensemble mean, and then averaging across all years T:

$$u=\frac{1}{T}\sum_{t=1}^{T}\frac{\max\left(\mathrm{peakSWE}_t\right)-\min\left(\mathrm{peakSWE}_t\right)}{\mathrm{mean}\left(\mathrm{peakSWE}_t\right)}.$$

We evaluate the temporal agreement in snowpack anomalies across datasets by calculating the Spearman's rank correlation coefficient of every combination of two products, generating a distribution of

$$\binom{15}{2}=105 \text{ coefficients for the CONUS,}$$
$$\binom{10}{2}=45 \text{ coefficients for the rest of North America, and}$$
$$\binom{9}{2}=36 \text{ coefficients for the rest of the Northern Hemisphere.}$$

We present the mean of that distribution to illustrate the expected value of the correlation. As a measure of significance, we hatch regions where fewer than half of all pairwise correlations are statistically significant ($p < 0.05$).

In the “Defining snow drought” section, we show the sensitivity of snow drought occurrence, characteristics, and trends to how snow droughts are defined. We illustrate this by applying six definitions of snow drought from the literature (Table A1) to basin-averaged SWE from the SNOTEL network of automated snow pillow sensors located in the Columbia River basin. Each SNOTEL site is weighted by the fraction of seasonally snow-covered area in the basin represented by the 100-m elevation band at which it is located. Only stations with 90% complete records since the start of WY 1980/81 are included (181 stations).

Figure 1a shows the timing and severity of the WY 2014/15 snow drought in the Columbia basin as determined by each metric. To assess how well these metrics agree with one another across time, we calculate percentiles of each metric on either daily (continuous SWE or SWEI) or yearly (peak or 1 April SWE) time scales relative to the entire WY 1980/81 to present record (Fig. 1b). Finally, to determine the consistency of definitions across the entire observational record, we create a binary snow drought classification based on each author’s metric and threshold (Fig. 1c). Snow drought definitions based on a single measurement (e.g., peak or 1 April SWE) are shown as covering the entire WY (Figs. 1b,c).

Our impacts analysis uses the U.S. Drought Monitor (USDM) weekly drought maps, which are jointly produced by the National Drought Mitigation Center at the University of Nebraska–Lincoln, the U.S. Department of Agriculture, and the National Oceanic and Atmospheric Administration. The USDM product is a synthesis of a wide array of numerical inputs, including multiple drought indices and observed and modeled estimates of soil moisture, streamflow, and vegetation health, with expert judgment and a network of local observers who provide information about how droughts are actually affecting people. Accordingly, we use it as an indicator of droughts that are truly impactful.

The areal extent of each level of drought is represented by a polygon on the USDM maps, so we rasterize these polygons to the same $0.5° \times 0.5°$ grid as the SWE ensemble. Because we are interested in causal linkages between snow droughts and warm-season droughts, we consider state changes in drought classes during the warm season (April–September). We consider a state change to be an onset, defined as a grid cell entering the warm season classified as not in drought and progressing to at least a moderate (D1) drought during those months, or an intensification, defined as an increase in the severity of drought during the warm season. Grid cells that enter the warm season in drought but do not experience an intensification are treated as not having a state change in drought.

To connect snow drought to its impacts, we adopt an approach similar to Livneh and Badger (2020) and calculate the change in forecast skill from the inclusion of snow drought as a predictor in a simple forecast model of basin-scale warm-season drought. As a measure of forecast skill, we use the equitable threat score (ETS, sometimes called the Gilbert skill score), a metric widely used in the weather forecasting community to evaluate forecasts of binary events, like the onset or intensification of warm-season drought. ETS is given by

$$\mathrm{ETS} = \frac{\mathrm{hits} - E}{\mathrm{hits} + \mathrm{misses} + \mathrm{false\ alarms} - E},$$

where a “hit” is a correctly predicted warm-season drought, a “miss” is an instance of a warm-season drought occurring when one was not predicted, and a “false alarm” is an instance of a predicted warm-season drought that does not occur. The E term accounts for the likelihood of “random hits” (i.e., it is easier to correctly predict droughts by chance in areas where they are frequent than places they are rare) (Gilbert 1884; Livneh and Badger 2020) and is given by

$$E = \frac{(\mathrm{hits} + \mathrm{misses})(\mathrm{hits} + \mathrm{false\ alarms})}{\mathrm{total}}.$$

To generate warm-season drought forecasts, we fit three sets of models for each basin. Our null model, representing the baseline drought forecast skill against which the snow drought models are compared, is given by

$$\Pr(\mathrm{WSD}) = \mathrm{logit}^{-1}\left(\beta_0 + \beta_1\mathbf{T} + \beta_2\mathbf{P}\right),$$

where Pr(WSD) is the probability of a warm-season drought onset or intensification and **T** and **P** are vectors of the average temperature anomaly and cumulative precipitation (as a percentage of the long-term median) in the first half of the WY (October–March) in the grid cells within the basin, respectively. Temperature and precipitation data are derived from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) reanalysis dataset (PRISM Climate Group 2004) and regridded to the 0.5° × 0.5° grid using bilinear interpolation. Because our outcome of interest, warm-season drought onset or intensification, is binary in nature, we employ a logistic regression model.

For each western U.S. basin (as delineated by the Global Runoff Data Centre) that contains SNOTEL sites (10 in total), we fit a separate model using a spatiotemporal panel of all grid cells contained within a basin (i.e., we use grid cell **T** and **P** to predict a state change in warm-season drought in that same grid cell). We estimate ETS using repeated fivefold cross validation. In this process, the data are randomly split into 5 equal-sized chunks of 3 years. A model is fit using 4 of these chunks (12 years), with the left-out 3 years of data being used to evaluate the forecast skill of the model. This is repeated for each of the 5 subsets, and the entire process is repeated 10 times, generating a distribution of 50 estimates of the out-of-sample forecast skill for the model in each.

We then consider the forecast skill gained from the inclusion of snow drought information based on a single dataset and definition. To that end, we specify another set of models that includes snow drought as a predictor:

$$\Pr\left(\mathrm{WSD}_{d,p}\right) = \mathrm{logit}^{-1}\left(\beta_0 + \beta_1\mathbf{T} + \beta_2\mathbf{P} + \beta_3\mathbf{SD}_{d,p}\right),$$

where $\mathbf{SD}_{d,p}$ is a binary classification of basin-scale snow drought as determined by definition d applied to data product p (see Fig. 3). Thus, in this specification, the predictors for warm-season drought are gridcell October–March average temperature and cumulative precipitation and a binary indicator of basin-scale snow drought. The use of basin-scale snow drought as a predictor of gridcell-level drought is designed to account for the fact that snow that accumulates far upstream in a basin can shape a downstream location's warm-season drought risk. We fit this model for all combinations of 16 datasets (the 15 members of the gridded ensemble plus SNOTEL) and 6 definitions.

Finally, to determine the value of leveraging the full 96-member ensemble and all observational and definitional uncertainty, we fit a third set of models that include the basin-scale ensemble mean that has been weighted by the skill of each constituent member:

$$\Pr(\mathrm{WSD}) = \mathrm{logit}^{-1}\left(\beta_0 + \beta_1\mathbf{T} + \beta_2\mathbf{P} + \beta_3\frac{1}{n_d n_p}\sum_d\sum_p \mathbf{W}_{d,p}\mathbf{SD}_{d,p}\right),$$

where n_d and n_p are the numbers of definitions and data products, respectively, and $\mathbf{W}_{d,p}$ is a weight proportional to the ETS of the forecast model using that particular data–definition combination. Because $\mathbf{SD}_{d,p}$ is binary, the ensemble mean in this context can be interpreted as the (weighted) proportion of ensemble members identifying a snow drought in a given WY.

In Fig. 3, we present the median of the distribution of 50 cross-validation estimates of ETS for both the null model (Fig. 3b) and the models using individual dataset–definition combinations (markers in Fig. 3a). For the model that includes the ensemble mean, we show the full distribution of skill estimates based on the cross validation, each plotted with a blue line with the same transparency; as more estimates overlap, the resulting region of higher opacity gives a confidence interval around the median estimate, which is indicated with a white line (Hsiang 2013; Mankin and Diffenbaugh 2015). To better visualize this confidence interval, we plot 1,000 samples from the probability distribution defined by the empirical cumulative distribution function of the 50 cross-validation estimates of forecast skill in each basin.

References

Abatzoglou, J. T., and A. P. Williams, 2016: Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. USA*, **113**, 11 770–11 775, https://doi.org/10.1073/pnas.1607171113.

Allen, C. D., D. D. Breshears, and N. G. McDowell, 2015: On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere*, **6** (8), 1–55, https://doi.org/10.1890/ES15-00203.1.

Annan, J. D., and J. C. Hargreaves, 2011: Understanding the CMIP3 multimodel ensemble. *J. Climate*, **24**, 4529–4538, https://doi.org/10.1175/2011JCLI3873.1.

——, and ——, 2017: On the meaning of independence in climate science. *Earth Syst. Dyn.*, **8**, 211–224, https://doi.org/10.5194/esd-8-211-2017.

Ashley, W. S., A. M. Haberlie, and V. A. Gensini, 2020: Reduced frequency and size of late-twenty-first-century snowstorms over North America. *Nat. Climate Change*, **10**, 539–544, https://doi.org/10.1038/s41558-020-0774-4.

Bach, A. F., G. Schrier, L. A. Melsen, A. M. G. K. Tank, and A. J. Teuling, 2018: Widespread and accelerated decrease of observed mean and extreme snow depth over Europe. *Geophys. Res. Lett.*, **45**, 122312–122319, https://doi.org/10.1029/2018GL079799.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438**, 303–309, https://doi.org/10.1038/nature04141.

Berg, N., and A. Hall, 2017: Anthropogenic warming impacts on California snowpack during drought. *Geophys. Res. Lett.*, **44**, 2511–2518, https://doi.org/10.1002/2016GL072104.

Berghuijs, W. R., R. A. Woods, and M. Hrachowitz, 2014: A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nat. Climate Change*, **4**, 583–586, https://doi.org/10.1038/nclimate2246.

Bintanja, R., and O. Andry, 2017: Towards a rain-dominated Arctic. *Nat. Climate Change*, **7**, 263–267, https://doi.org/10.1038/nclimate3240.

Brown, R., and B. Brasnett, 2010: Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data, version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 17 February 2021, https://doi.org/10.5067/W9FOYWH0EQZ3.

Broxton, P., X. Zeng, and N. Dawson, 2019: Daily 4 km gridded SWE and snow depth from assimilated in-situ and modeled data over the conterminous US, version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 14 December 2020, https://doi.org/10.5067/0GGPB220EX6A.

Cayan, D. R., S. A. Kammerdiener, M. D. Dettinger, J. M. Caprio, and D. H. Peterson, 2001: Changes in the onset of spring in the western United States. *Bull. Amer. Meteor. Soc.*, **82**, 399–416, https://doi.org/10.1175/1520-0477(2001)082<0399:CITOOS>2.3.CO;2.

Cook, B. I., A. P. Williams, J. S. Mankin, R. Seager, J. E. Smerdon, and D. Singh, 2018: Revisiting the leading drivers of Pacific coastal drought variability in the contiguous United States. *J. Climate*, **31**, 25–43, https://doi.org/10.1175/JCLI-D-17-0172.1.

Cooper, M. G., A. W. Nolin, and M. Safeeq, 2016: Testing the recent snow drought as an analog for climate warming sensitivity of Cascades snowpacks. *Environ. Res. Lett.*, **11**, 084009, https://doi.org/10.1088/1748-9326/11/8/084009.

Dierauer, J. R., D. M. Allen, and P. H. Whitfield, 2019: Snow drought risk and susceptibility in the western United States and southwestern Canada. *Water Resour. Res.*, **55**, 3076–3091, https://doi.org/10.1029/2018WR023229.

Dietz, A. J., C. Kuenzer, U. Gessner, and S. Dech, 2012: Remote sensing of snow – A review of available methods. *Int. J. Remote Sens.*, **33**, 4094–4134, https://doi.org/10.1080/01431161.2011.640964.

Diffenbaugh, N. S., M. Scherer, and M. Ashfaq, 2013: Response of snow-dependent hydrologic extremes to continued global warming. *Nat. Climate Change*, **3**, 379–384, https://doi.org/10.1038/nclimate1732.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gilbert, G. K., 1884: Finley's Tornado predictions. *Amer. Meteor. J.*, **1**, 166–172.

GRDC, 2020: *Major River Basins of the World.* 2nd ed. Global Runoff Data Centre, www.bafg.de/GRDC/EN/02_srvcs/22_gslrs/221_MRB/riverbasins_node.html.

Gringorten, I. I., 1963: A plotting rule for extreme probability paper. *J. Geophys. Res.*, **68**, 813–814, https://doi.org/10.1029/JZ068i003p00813.

Groisman, P. Ya., T. R. Karl, and R. W. Knight, 1994: Observed impact of snow cover on the heat balance and the rise of continental spring temperatures. *Science*, **263**, 198–200, https://doi.org/10.1126/science.263.5144.198.

Hall, A., X. Qu, and J. D. Neelin, 2008: Improving predictions of summer climate change in the United States. *Geophys. Res. Lett.*, **35**, L01702, https://doi.org/10.1029/2007GL032012.

Harley, G. L., R. S. Maxwell, B. A. Black, and M. F. Bekker, 2020: A multi-century, tree-ring-derived perspective of the North Cascades (USA) 2014–2016 snow drought. *Climatic Change*, **162**, 127–143, https://doi.org/10.1007/s10584-020-02719-0.

Harpold, A. A., and N. P. Molotch, 2015: Sensitivity of soil water availability to changing snowmelt timing in the western U.S. *Geophys. Res. Lett.*, **42**, 8011–8020, https://doi.org/10.1002/2015GL065855.

——, M. D. Dettinger, and S. Rajagopal, 2017: Defining snow drought and why it matters. *Eos*, **98**, https://doi.org/10.1029/2017EO068775.

Hartmann, H., and Coauthors, 2018: Research frontiers for improving our understanding of drought-induced tree and forest mortality. *New Phytol.*, **218**, 15–28, https://doi.org/10.1111/nph.15048.

Hatchett, B. J., and D. J. McEvoy, 2017: Exploring the origins of snow drought in the northern Sierra Nevada, California. *Earth Interact.*, **22**, https://doi.org/10.1175/EI-D-17-0027.1.

Henderson, G. R., Y. Peings, J. C. Furtado, and P. J. Kushner, 2018: Snow–atmosphere coupling in the Northern Hemisphere. *Nat. Climate Change*, **8**, 954–963, https://doi.org/10.1038/s41558-018-0295-6.

Hodgkins, G. A., and R. W. Dudley, 2006: Changes in late-winter snowpack depth, water equivalent, and density in Maine, 1926–2004. *Hydrol. Processes*, **20**, 741–751, https://doi.org/10.1002/hyp.6111.

Hsiang, S., 2013: Visually-weighted regression. SSRN Scholarly Paper ID 2265501, Social Science Research Network, 10 pp., https://doi.org/10.2139/ssrn.2265501.

Huning, L. S., and A. AghaKouchak, 2018: Mountain snowpack response to different levels of warming. *Proc. Natl. Acad. Sci. USA*, **115**, 102932–102937, https://doi.org/10.1073/pnas.1805953115.

——, and ——, 2020: Global snow drought hot spots and characteristics. *Proc. Natl. Acad. Sci. USA*, **117**, 192753–192759, https://doi.org/10.1073/pnas.1915921117.

Kapnick, S. B., and T. L. Delworth, 2013: Controls of global snow under a changed climate. *J. Climate*, **26**, 5537–5562, https://doi.org/10.1175/JCLI-D-12-00528.1.

Kim, R. S., and Coauthors, 2021: Snow Ensemble Uncertainty Project (SEUP): Quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling. *Cryosphere*, **15**, 771–791, https://doi.org/10.5194/tc-15-771-2021.

Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, https://doi.org/10.2151/jmsj.2015-001.

Kolstad, E. W., 2017: Causal pathways for temperature predictability from snow depth. *J. Climate*, **30**, 9651–9663, https://doi.org/10.1175/JCLI-D-17-0280.1.

Koster, R. D., S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier, and R. H. Reichle, 2010: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat. Geosci.*, **3**, 613–616, https://doi.org/10.1038/ngeo944.

Krasting, J. P., A. J. Broccoli, K. W. Dixon, and J. R. Lanzante, 2013: Future changes in Northern Hemisphere snowfall. *J. Climate*, **26**, 7813–7828, https://doi.org/10.1175/JCLI-D-12-00832.1.

Livneh, B., and A. M. Badger, 2020: Drought less predictable under declining future snowpack. *Nat. Climate Change*, **10**, 452–458, https://doi.org/10.1038/s41558-020-0754-8.

Luojus, K., J. Pulliainen, M. Takala, J. Lemmetyinen, and M. Moisander, 2020: GlobSnow v3.0 snow water equivalent (SWE). PANGAEA, accessed 14 December 2020, https://doi.org/10.1594/PANGAEA.911944.

Mankin, J. S., and N. S. Diffenbaugh, 2015: Influence of temperature and precipitation variability on near-term snow trends. *Climate Dyn.*, **45**, 1099–1116, https://doi.org/10.1007/s00382-014-2357-4.

——, D. Viviroli, D. Singh, A. Y. Hoekstra, and N. S. Diffenbaugh, 2015: The potential for snow to supply human water demand in the present and future. *Environ. Res. Lett.*, **10**, 114016, https://doi.org/10.1088/1748-9326/10/11/114016.

——, R. Seager, J. E. Smerdon, B. I. Cook, and A. P. Williams, 2019: Mid-latitude freshwater availability reduced by projected vegetation responses to climate change. *Nat. Geosci.*, **12**, 983–988, https://doi.org/10.1038/s41561-019-0480-x.

Marshall, A. M., J. T. Abatzoglou, T. E. Link, and C. J. Tennant, 2019: Projected changes in interannual variability of peak snowpack amount and timing in the western United States. *Geophys. Res. Lett.*, **46**, 8882–8892, https://doi.org/10.1029/2019GL083770.

Marty, C., A.-M. Tilg, and T. Jonas, 2017: Recent evidence of large-scale receding snow water equivalents in the European Alps. *J. Hydrometeor.*, **18**, 1021–1031, https://doi.org/10.1175/JHM-D-16-0188.1.

McCabe, G. J., and M. D. Dettinger, 2002: Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate. *J. Hydrometeor.*, **3**, 13–25, https://doi.org/10.1175/1525-7541(2002)003<0013:PMAPOY>2.0.CO;2.

Meira Neto, A. A., G.-Y. Niu, T. Roy, S. Tyler, and P. A. Troch, 2020: Interactions between snow cover and evaporation lead to higher sensitivity of streamflow to temperature. *Commun. Earth Environ.*, **1**, 56, https://doi.org/10.1038/s43247-020-00056-9.

Meromy, L., N. P. Molotch, T. E. Link, S. R. Fassnacht, and R. Rice, 2013: Subgrid variability of snow water equivalent at operational snow stations in the western USA. *Hydrol. Processes*, **27**, 2383–2400, https://doi.org/10.1002/hyp.9355.

Molotch, N. P., and R. C. Bales, 2006: SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence. *Hydrol. Processes*, **20**, 723–739, https://doi.org/10.1002/hyp.6128.

——, P. D. Brooks, S. P. Burns, M. Litvak, R. K. Monson, J. R. McConnell, and K. Musselman, 2009: Ecohydrological controls on snowmelt partitioning in mixed-conifer sub-alpine forests. *Ecohydrology*, **2**, 129–142, https://doi.org/10.1002/eco.48.

Mortimer, C., L. Mudryk, C. Derksen, K. Luojus, R. Brown, R. Kelly, and M. Tedesco, 2020: Evaluation of long-term Northern Hemisphere snow water equivalent products. *Cryosphere*, **14**, 1579–1594, https://doi.org/10.5194/tc-14-1579-2020.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining mountain snowpack in western North America. *Bull. Amer. Meteor. Soc.*, **86**, 39–50, https://doi.org/10.1175/BAMS-86-1-39.

——, and Coauthors, 2016: Perspectives on the causes of exceptionally low 2015 snowpack in the western United States. *Geophys. Res. Lett.*, **43**, 10 980–10 988, https://doi.org/10.1002/2016GL069965.

——, S. Li, D. P. Lettenmaier, M. Xiao, and R. Engel, 2018: Dramatic declines in snowpack in the western US. *npj Climate Atmos. Sci.*, **1**, 2, https://doi.org/10.1038/s41612-018-0012-1.

Mudryk, L. R., C. Derksen, P. J. Kushner, and R. Brown, 2015: Characterization of Northern Hemisphere snow water equivalent datasets, 1981–2010. *J. Climate*, **28**, 8037–8051, https://doi.org/10.1175/JCLI-D-15-0229.1.

Musselman, K. N., N. Addor, J. A. Vano, and N. P. Molotch, 2021: Winter melt trends portend widespread declines in snow water resources. *Nat. Climate Change*, **11**, 418–424, https://doi.org/10.1038/s41558-021-01014-9.

National Operational Hydrologic Remote Sensing Center, 2004: Snow Data Assimilation System (SNODAS) data products at NSIDC, version 1. National Snow and Ice Data Center, accessed 23 November 2020, https://doi.org/10.7265/N5TB14TC.

Pagano, T., D. Garen, and S. Sorooshian, 2004: Evaluation of official western U.S. seasonal water supply outlooks, 1922–2002. *J. Hydrometeor.*, **5**, 896–909, https://doi.org/10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2.

Pierce, D. W., and D. R. Cayan, 2013: The uneven response of different snow measures to human-induced climate warming. *J. Climate*, **26**, 4148–4167, https://doi.org/10.1175/JCLI-D-12-00534.1.

——, and Coauthors, 2008: Attribution of declining western U.S. snowpack to human effects. *J. Climate*, **21**, 6425–6444, https://doi.org/10.1175/2008JCLI2405.1.

Potopová, V., C. Boroneanț, M. Možný, and J. Soukup, 2016: Driving role of snow cover on soil moisture and drought development during the growing season in the Czech Republic. *Int. J. Climatol.*, **36**, 3741–3758, https://doi.org/10.1002/joc.4588.

PRISM Climate Group, 2004: PRISM Climate Data. Oregon State University, http://prism.oregonstate.edu.

Pulliainen, J., and Coauthors, 2017: Early snowmelt significantly enhances boreal springtime carbon uptake. *Proc. Natl. Acad. Sci. USA*, **114**, 11 081–11 086, https://doi.org/10.1073/pnas.1707889114.

——, and Coauthors, 2020: Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. *Nature*, **581**, 294–298, https://doi.org/10.1038/s41586-020-2258-0.

Qin, Y., and Coauthors, 2020: Agricultural risks from changing snowmelt. *Nat. Climate Change*, **10**, 459–465, https://doi.org/10.1038/s41558-020-0746-8.

Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, https://doi.org/10.1175/BAMS-85-3-381.

Sabater, J. M., 2019: ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed 27 August 2020, https://doi.org/10.24381/cds.e2161bac.

Seneviratne, S. I., D. Lüthi, M. Litschi, and C. Schär, 2006: Land–atmosphere coupling and climate change in Europe. *Nature*, **443**, 205–209, https://doi.org/10.1038/nature05095.

Siler, N., C. Proistosescu, and S. Po-Chedley, 2019: Natural variability has slowed the decline in western U.S. snowpack since the 1980s. *Geophys. Res. Lett.*, **46**, 346–355, https://doi.org/10.1029/2018GL081080.

Smith, A. B., 2020: U.S. billion-dollar weather and climate disasters, 1980–present (NCEI Accession 0209268). NOAA/NCEI, accessed 14 December 2020, https://doi.org/10.25921/stkw-7w73.

Stephens, S. L., and Coauthors, 2018: Drought, tree mortality, and wildfire in forests adapted to frequent fire. *BioScience*, **68**, 77–88, https://doi.org/10.1093/biosci/bix146.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward earlier streamflow timing across western North America. *J. Climate*, **18**, 1136–1155, https://doi.org/10.1175/JCLI3321.1.

Tedesco, M., and J. Jeyaratnam, 2019: AMSR-E/AMSR2 Unified L3 Global Daily 25 km EASE-Grid Snow Water Equivalent, version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 11 November 2020, https://doi.org/10.5067/8AE2ILXB5SM6.

Thornton, M., R. Shrestha, Y. Wei, P. Thornton, S. Kao, and B. Wilson, 2020: Daymet: Daily Surface Weather Data on a 1-km Grid for North America, version 4. ORNL Distributed Active Archive Center, accessed 29 December 2020, https://doi.org/10.3334/ORNLDAAC/1840.

Trujillo, E., N. P. Molotch, M. L. Goulden, A. E. Kelly, and R. C. Bales, 2012: Elevation-dependent influence of snow accumulation on forest greening. *Nat. Geosci.*, **5**, 705–709, https://doi.org/10.1038/ngeo1571.

Ukkola, A. M., I. C. Prentice, T. F. Keenan, A. I. J. M. van Dijk, N. R. Viney, R. B. Myneni, and J. Bi, 2016: Reduced streamflow in water-stressed climates consistent with CO_2 effects on vegetation. *Nat. Climate Change*, **6**, 75–78, https://doi.org/10.1038/nclimate2831.

van den Hurk, B., and Coauthors, 2016: LS3MIP (v1.0) contribution to CMIP6: The Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome. *Geosci. Model Dev.*, **9**, 2809–2832, https://doi.org/10.5194/gmd-9-2809-2016.

Viviroli, D., H. H. Dürr, B. Messerli, M. Meybeck, and R. Weingartner, 2007: Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resour. Res.*, **43**, W07447, https://doi.org/10.1029/2006WR005653.

Westerling, A. L., 2016: Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Philos. Trans. Roy. Soc. London*, **B371**, 20150178, https://doi.org/10.1098/rstb.2015.0178.

——, H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, 2006: Warming and earlier spring increase western U.S. forest wildfire activity. *Science*, **313**, 940–943, https://doi.org/10.1126/science.1128834.

Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, https://doi.org/10.1029/2011JD016048.

Zeng, X., P. Broxton, and N. Dawson, 2018: Snowpack change from 1982 to 2016 over conterminous United States. *Geophys. Res. Lett.*, **45**, 12 940–12 947, https://doi.org/10.1029/2018GL079621.

Zhang, T., T. Wang, G. Krinner, X. Wang, T. Gasser, S. Peng, S. Piao, and T. Yao, 2019: The weakening relationship between Eurasian spring snow cover and Indian summer monsoon rainfall. *Sci. Adv.*, **5**, eaau8932, https://doi.org/10.1126/sciadv.aau8932.