https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Algorithmically Detected Rain-on-Snow Flood Events in Different Climate Datasets: A Case Study of the Susquehanna River Basin

Colin M. Zarzycki[1], Benjamin D. Ascher[1,*], Alan M. Rhoades[2], and Rachel R. McCrary[3]

[1]Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA, USA
[2]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3]NSF, National Center for Atmospheric Research, Research Applications Laboratory, Boulder, CO, USA
[*]now at Colorado State University

Correspondence: Colin M. Zarzycki (czarzycki@psu.edu)

Abstract. Rain-on-snow (RoS) events in regions of ephemeral snowpack – such as the northeastern United States – can be key drivers of cool-season flooding. We describe an automated algorithm for detecting basin-scale RoS events in gridded climate data by generating an area-averaged time-series and then searching for periods of concurrent precipitation, surface runoff, and snowmelt exceeding pre-defined thresholds. When evaluated using historical data over the Susquehanna River Basin (SRB), the technique credibly finds RoS events in published literature and flags events that are followed by anomalously high streamflow as measured by gage data along the river. When comparing four different datasets representing the same 21-year period, we find large differences in RoS event magnitude and frequency, primarily driven by differences in estimated surface runoff and snowmelt. Using dataset-specific thresholds improves agreement between datasets but does not account for all discrepancies. We show that factors such as meteorological forcing and coupling frequency as well as choice of land surface model play roles in how data products capture these compound extremes and suggest care is to be taken when climate datasets are used by stakeholders for operational decision-making.

1 Introduction

Rain-on-snow (RoS) events have been increasingly studied over the past few decades, yet such research is overwhelmingly focused on mountainous regions with well-defined seasonal snowpacks (Singh et al., 1997; McCabe et al., 2007; Wayand et al., 2015; Sterle et al., 2019; Musselman et al., 2018; Poschlod et al., 2020; Hatchett, 2021; Siirila-Woodburn et al., 2021b; Heggli et al., 2022; Yu et al., 2022; Brandt et al., 2022; Maina and Kumar, 2023; Haleakala et al., 2023). Correspondingly, there has been less focus in areas with more ephemeral snow cover, such as the northeastern United States (US), even though RoS events, a flavor of compound extreme (AghaKouchak et al., 2020), produce many cases of 'slow-rise' flooding – floods generally occurring more than six hours after the onset of the meteorological driver (Dougherty et al., 2021). Climatologically, RoS events in the northeastern US peak in late winter and spring (Ashley and Ashley, 2008; Villarini and Smith, 2010; Dougherty and Rasmussen, 2019; Wachowicz et al., 2020) and are key drivers of flooding in New England and the Atlantic side of Canada (Collins et al., 2014). Synoptic case study analyses of recent RoS events in the mid-Atlantic highlight inland-running extratropical cyclones that advect warm, moist air into the region as key dynamical drivers (Grote, 2021; Suriano et al., 2023).

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Rapid snow ablation in ephemeral snow regions has been shown to have a strong correlation with increases in basin streamflow in the days following snowmelt (Suriano et al., 2020, 2023).

While climatological studies are critically important, event-level analysis has become increasingly valuable when communicating climate risks (Shepherd et al., 2018). One river basin in the northeastern US that has historically dealt with RoS events is the Susquehanna River Basin (SRB) – a basin home to more than four million people (Leathers et al., 2008) and one considered climatologically flood-prone due to the wide variety of weather phenomena that occur within the region (Perry, 2000). The most consequential non-tropical-cyclone flood in recent SRB history was a RoS flood that occurred in January 1996, resulting in ∼\$1.5 billion in damages and 30 fatalities (Leathers et al., 1998). Significant events such as this are frequently used by stakeholders as a point of reference for real-time forecasts and long-term planning (George and Mudelsee, 2019). Other evidence, such as the Great Flood of 1936 (another RoS event) and the fact that sediment records indicate prehistoric periods of high flood activity are associated with negative phases of the North Atlantic Oscillation (NAO) (which drive positive snowpack anomalies in the northeastern US, (Hartley and Keables, 1998)) underscore the importance of these events to regional hydrology (Toomey et al., 2019).

While studying historical events is important for planning purposes, and historical reforecasts using imposed warming approaches can provide clues as to how similar events may unfold in the future (e.g., Pettett and Zarzycki (2023)), it is difficult for such approaches to provide risk quantification from a frequency-of-occurrence perspective. For these climatological evaluations, it is desirable to be able to identify such events in both observational datasets and model simulations, including future climate projections. Unfortunately, assessing extreme, compound, and discrete events in climate data sets is a complex challenge, particularly because such datasets are not necessarily developed specifically for this purpose (Angélil et al., 2017; Parker, 2020) and there is commonly a lack of observational reference datasets to quantify their fidelity. Gridded datasets (spatiotemporally continuous data provided on a regular latitude-longitude mesh) of the historical record are frequently used to assess hydrometeorological extremes in locations with poor or non-existent station observations and climate analyses requiring multiple complex variables in addition to evaluating model sensitivity and performance. Further, future changes in RoS event frequency and character due to climate change are projected via the use of free-running models, which operate on numerical grids and don't have an *a priori* record to compare to, making the use of an automated heuristic a requirement.

Therefore, to extract information regarding compound hydrometeorological extremes, such as RoS events, and their corresponding statistics, algorithmic techniques that objectively analyze datasets without manual intervention are desirable. Here, we demonstrate a technique for generating a RoS event database at the basin scale for arbitrary gridded datasets and intercompare key decision-relevant differences (e.g., flood frequency) across four climate data products. We choose to focus on the SRB based on its proximity to major population centers, existing evidence for increasing flood hazards and exposure in the basin (Sharma et al., 2021), and the aforementioned 1996 extreme event serving as a benchmark for the mid-Atlantic US, although the technique described can be applied to any geographically defined basin with properly specified thresholds.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

2 Methods

2.1 Datasets

We evaluate RoS events within three widely-used climate datasets and in one state-of-the-art Earth system model (ESM) nudged towards an atmospheric reanalysis. All four datasets seek to reproduce observed conditions, although each uses a distinct methodology to do so. First, we investigate the dataset described in Livneh et al. (2015) (hereafter, L15). L15 is a widely-used 1/16° hydrometeorological dataset covering most of North America. Meteorological data provided by L15 consists of daily precipitation, temperature (maximum and minimum), and wind speed at each location (Henn et al., 2018). This data is then temporally interpolated to obtain subdaily estimates (Bohn et al., 2013), which are then used to drive the Variable Infiltration Capacity (VIC, Liang et al. (1994)) land surface model (LSM) to produce hydrometeorological outputs. Next, we investigate the 1/8° North American Land Data Assimilation System (NLDAS-VIC4.0.5) described in Xia et al. (2012). NLDAS is driven by offline atmospheric forcing derived from the North American Regional Reanalysis, with adjustments made to some variables based on observations. NLDAS uses a combination of daily observations and radar data to produce hourly estimates of precipitation. While there are multiple NLDAS-2 LSMs that produce hydrologic output variables, we analyze only the NLDAS-VIC4.0.5 as it is the most methodologically consistent with L15. We also analyze a ∼0.5° global reanalysis (JRA-55, Kobayashi et al. (2015)), generated by running a prognostic ESM while continually assimilating observational data during integration. Global reanalyses serve as a bridge between in-situ, but spatio-temporally unstructured, observational data and free-running climate models (Parker, 2016). Finally, a ∼1° nudged version of the U.S. Department of Energy (DOE) Energy Exascale Earth System Model (E3SM, Golaz et al. (2022)) model is analyzed (Zhang et al., 2022). Assessing this dataset provides insight into the capability of ESMs used in climate assessments (e.g., the Coupled Model Intercomparison Project Phase 6 (CMIP6), Eyring et al. (2016)) to capture observed hydrometeorological extreme events. E3SM performs close to the CMIP6 ensemble mean across several measures of skill (Fasullo, 2020). To constrain the large-scale meteorology in E3SM to match observed conditions, the E3SM simulation is nudged using 6-hourly fields from the ERA5 reanalysis (Hersbach et al., 2020) generated using the technique contained in the Betacast toolkit, initially outlined in Zarzycki and Jablonowski (2015). This nudging acts as a crude assimilation technique and is only applied in the free atmosphere (above nominally 850 hPa, or 1 kilometer). The zonal and meridional winds are nudged to ERA5 analysis with a relaxation timescale τ = 3 hours and the temperature is nudged with τ = 24 hours. No nudging is performed on the surface pressure or moisture fields. The simulation reproduces observed patterns of 500 hPa geopotential height and sea level pressure while simulating grid-scale processes relatively untethered by the nudging reanalysis product (Sun et al., 2019). E3SM is run with a spectral element dynamical core on an unstructured cubed-sphere mesh ($n_e 30 n_p 4$) and simulation output is remapped to a 1°x1° rectilinear grid using higher-order methods (Hill et al., 2004). All model settings and tunings are left as the default contained in the commit used here (f9cbe57).

2.2 Defining basin-scale events

To identify RoS events, three hydrometeorological variables are extracted over the SRB. These include precipitation (PRECIP), snow water equivalent (SWE), and surface water runoff (ROF). The 24-hour change in SWE from the previous day (dSWE) is calculated via a simple backward difference. All data is standardized to daily average values (00Z to 00Z) by temporally averaging any sub-daily (e.g., hourly) data at each grid cell. Grid cells that have at least 50% of their area enclosed by the boundary of the SRB are kept, while those exterior to the SRB are set to missing values. The resulting domains for each product can be seen in Fig. 1. A basin-wide time series is then constructed by spatially averaging fields for each day and smoothing the resulting time series using a moving average to reduce day-to-day noise. We choose a 5-day window based on mid-latitude synoptic timescales (Holton, 2004), although other window durations did not materially impact these findings (not shown).

This results in a one-dimensional time series with a single value for each calendar day that represents the SRB basin-averaged conditions for each data product. RoS days are defined by flagging periods of negative dSWE and positive ROF that both exceed specified thresholds (t) on the same days. Contiguous days are considered to be part of a single 'event' such that discrete events can last for different durations. We test two methods of thresholding; one uses fixed thresholds across all four datasets (FIXED) and the other defines dataset-specific thresholds by those exceeding 95% of all daily values (RELATIVE). For FIXED, we require an average dSWE of -1.4 mm/day (t_{dSWE}) and ROF of 1.4 mm/day (t_{ROF}) averaged across the SRB based on a manual sensitivity analysis. Thresholds in the RELATIVE configuration are calculated independently for each dataset by setting the threshold to the 95% value of its own 1985-2005 time series. To enforce a criterion that precipitation occurs during at least some portion of the event, we require an average PRECIP of 2 mm/day (t_{PRECIP}), which is kept the same in both the FIXED and RELATIVE configurations. We note that our findings are actually largely insensitive to the magnitude of the PRECIP threshold (or even its inclusion), implying the vast majority of events with high dSWE and ROF are also associated with non-zero PRECIP. This supports the notion that similar synoptic meteorological patterns (Grote, 2021) and additional liquid input to the surface (i.e., runoff being a combination of snowmelt and water flux from the atmosphere) lead to the majority of RoS floods being associated with, at least, some precipitation versus precipitation alone being a primary driver of such events. It also concurs with Suriano et al. (2023), who found that, while RoS days contributed disproportionately to extreme snow ablation events, snowmelt also occurred in a myriad of non-precipitating patterns, including high-pressure overhead and under northwesterly/westerly large-scale flow. We argue this also corroborates the finding that actual heat transfer between the liquid rain and the surface of the snowpack is rather small and explains only a small fraction of the observed snowmelt (Moore and Owens, 1984).

For context, our definition here of RoS is somewhat arbitrary. Most studies enforce some fixed combination of precipitation and snowpack threshold. Ye et al. (2008) classified events as liquid precipitation falling onto at least 10 mm of existing snow-pack in a gridbox. Musselman et al. (2018) define 'RoS days' where at least 10 mm of precipitation falls on at least 10 mm of snowpack in a given day. This technique was adopted by López-Moreno et al. (2021) and Maina and Kumar (2023). A similar strategy, but applying higher thresholds (20 mm precipitation, 250 mm snowpack), was used by Würzer et al. (2016). Likewise,

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Hotovy et al. (2023) required 10 mm of existing SWE and greater than 0°C surface temperatures to enforce collocated daily precipitation (of at least 5 mm) to fall in liquid form. Instead of a SWE metric, others have applied snow cover fraction or some other snow/no-snow classification in addition to a precipitation threshold (Mazurkiewicz et al., 2008; Pradhanang et al., 2013; Cohen et al., 2015). Others include some measure of snowmelt, such as Freudiger et al. (2014) and Li et al. (2019) who required 3 mm of rainfall, 10 mm SWE, and a 20% reduction in the snowpack (implied dSWE <= -20 mm/day). Suriano (2022) similarly required a 10 mm daily snow depth decrease that occurred in combination with above-freezing temperatures and non-zero precipitation. However, relative strategies have also been proposed, such as the 98% threshold used for covarying extremes in the compound event analysis of Poschlod et al. (2020). Other studies simply require the joint occurrence of precipitation and snowmelt in a given period (McCabe et al., 2007; Surfleet and Tullos, 2013; Collins et al., 2014; Guan et al., 2016; Jeong and Sushama, 2018). Wachowicz et al. (2020) used the same gridded data and explored four different definitions derived from some of the above and found high-level agreement.

This brief literature review is not meant to be considered exhaustive, but rather to highlight that both RoS evaluation techniques in this manuscript fall within the scope of existing strategies. It also emphasizes that extreme RoS events in mountainous regions with more seasonal snowpack or those using gridpoint values versus basin-integrated metrics may require different thresholds. With the caveat that much of the previously cited RoS work has focused on regions with less ephemeral snowpack (e.g., western U.S. mountains), the algorithm discussed here falls within the envelope of previously published results both from a heuristic perspective and also with respect to our defined thresholds.

Lastly, to show all datasets adequately represent the same period, Pearson correlations of basin-wide statistics (Table 1) were statistically significant with a two-sided t-test between all permutations of daily time series. This confirms that all datasets fundamentally represent the same meteorology and land surface evolution as processed here and can therefore be directly compared to one another.

3 Results

3.1 Climatology

We first investigate the mean climatology of relevant quantities over the SRB to provide some context into each data product's baseline. The average cool-season (November-April, inclusive) distributions of PRECIP, ROF, SWE, and dSWE are shown in Fig. 1. The higher resolution of L15 and NLDAS is evident from the added structure in the mean fields of the first two columns. Mean PRECIP (Fig. 1a-d) is higher in JRA and E3SM when compared to the L15 and NLDAS. This may be due to factors such as atmospheric model biases or the inclusion of rain gauge data in L15 and NLDAS, although it has been shown that significant spread exists in historical gridded climate data products, even for more commonly observed variables such as precipitation (Gutmann et al., 2012; Livneh et al., 2014; Henn et al., 2018). ROF (Fig. 1e-h) and SWE (Fig. 1i-l) climatologies differ between the data products. Notably, L15 produces mean ROF values that are less than half of the climatologies of each of the other three datasets. It is well-known that simulated ROF from different hydrologic models can be extremely variable, with regional differences between products reaching an order of magnitude (Gudmundsson et al., 2012; Sood and Smakhtin,

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Table 1. Pearson correlations between data products for daily timeseries of precipitation (top), surface runoff (middle), and 24-hour change in snow water equivalent (bottom) over the study period. All values are correlated significantly at greater than the 99.9% confidence level using a two-sided t-test, indicating that all timeseries represent the same historical period and include relevant day-to-day variations over the SRB.

PRECIP				
	JRA	L15	NLDAS	E3SM
JRA	1.00	0.76	0.89	0.75
L15	0.76	1.00	0.81	0.93
NLDAS	0.89	0.81	1.00	0.77
E3SM	0.75	0.93	0.77	1.00

ROF				
	JRA	L15	NLDAS	E3SM
JRA	1.00	0.89	0.88	0.83
L15	0.89	1.00	0.85	0.74
NLDAS	0.88	0.85	1.00	0.82
E3SM	0.83	0.74	0.82	1.00

dSWE				
	JRA	L15	NLDAS	E3SM
JRA	1.00	0.34	0.46	0.47
L15	0.34	1.00	0.79	0.66
NLDAS	0.46	0.79	1.00	0.71
E3SM	0.47	0.66	0.71	1.00

2015; Beck et al., 2017). NLDAS produces less SWE climatologically, <20% of the SWE produced by L15 or E3SM, and even less than the coarser JRA product. Previous work has also shown that SWE estimates can vary greatly across datasets (Lundquist et al., 2015; Rhoades et al., 2018), particularly over the ephemeral snow area of the northeastern U.S. (McCrary et al., 2017, 2022). dSWE climatology is shown in Fig. 1m-p. When calculating the mean, all accumulation (or zero change) days are ignored to isolate only days where snow loss occurred. Here, NLDAS also exhibits the lowest magnitude of dSWE, although we speculate this is at least partly due to the shallower mean snowpack. The largest climatological dSWE magnitude is found in JRA, even though the SWE does not contain the largest depths, implying the snowpack is more variable in JRA and may be prone to more rapid snow loss (from a dSWE per unit time perspective).

We emphasize that, from a physical standpoint, differences in snowfall and snowmelt timing (Rauscher et al., 2008; McCabe and Clark, 2005), snow properties (Brown et al., 2006), temperature and permeability of the soil (Niu and Yang, 2006), precipitation type partitioning (Knowles et al., 2006), and evapotranspiration (Zheng et al., 2019) all can lead to differences in how these quantities are simulated. We speculate that these mechanisms are playing important roles in the differing mean

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Table 2. Statistics of RoS events over the SRB using FIXED thresholds (top) and RELATIVE thresholds (bottom). t_{dSWE}, t_{ROF}, and t_{PRECIP} represent the thresholds used for snow water equivalent loss, surface runoff, and precipitation, respectively (mm/day). Events represent the number of RoS events flagged over the 1980-2005 period. Duration is the average number of consecutive days an RoS event lasts. dSWE$_a$, ROF$_a$, and PRECIP$_a$ represent the average snow loss, amount of runoff rate, and amount of precipitation rate (mm/day) per event by calculating the mean daily value for each individual event and then averaging those.

	t_{dSWE}	t_{ROF}	t_{PRECIP}	Events	Duration	dSWE$_a$	ROF$_a$	PRECIP$_a$
	mm day^{-1}	*mm day*$^{-1}$	*mm day*$^{-1}$	*#*	*days*	*mm day*$^{-1}$	*mm day*$^{-1}$	*mm day*$^{-1}$
FIXED								
L15	-1.4	1.4	2.0	6	3.2	-5.3	1.5	9.5
NLDAS	-1.4	1.4	2.0	16	2.9	-5.2	2.1	6.3
JRA	-1.4	1.4	2.0	48	3.5	-4.4	2.6	5.3
E3SM	-1.4	1.4	2.0	58	4.1	-5.2	2.4	6.5
RELATIVE								
L15	-1.5	0.6	2.0	20	5.2	-5.3	1.1	6.2
NLDAS	-0.8	1.4	2.0	20	2.8	-4.2	2.1	7.2
JRA	-1.9	1.6	2.0	40	3.2	-4.1	2.5	5.1
E3SM	-2.2	1.8	2.0	41	4.1	-8.0	2.9	6.9

climatologies but performing a fully-detailed water budget analysis for each of the land surface models leveraged by these datasets is beyond the scope of this analysis.

3.2 Flagged event statistics

Turning to the algorithmically flagged RoS events, Table 2 shows summary statistics for all four data products. Focusing on the FIXED thresholds, the total number of discrete events that occurred in the SRB over the 21-year study period varies by an order of magnitude, from 6 events in L15 to 58 in E3SM. Interestingly, large differences don't necessarily appear when considering the mean event-averaged dSWE, ROF, or PRECIP (last three columns). In fact, when a RoS event occurs, L15 has the largest PRECIP$_a$ and largest dSWE$_a$, although the smallest ROF$_a$.

When using the RELATIVE thresholds, the event frequencies agree better, with only a factor of two difference between L15/NLDAS and JRA/E3SM. The number of events in L15 increases because the ROF threshold (95th percentile of daily values) is reduced from 1.4 to 0.6 mm/day. Similarly, the dSWE threshold magnitude is reduced from -1.4 mm/day to -0.8 mm/day in NLDAS. Conversely, fewer events were classified in E3SM, due to increases in both the required ROF and dSWE thresholds applied to the daily climatology. However, even when accounting for the baseline climatological differences of the data products by thresholding on percentiles, rather than absolute magnitudes, differences still are evident in all metrics.

The differences across gridded climate data products are also shown in Fig. 2 (FIXED) and in Fig. 3 (RELATIVE). The sign of dSWE is flipped here for ease of visualization with the other metrics (i.e., positive dSWE in both figures denotes snowpack).

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

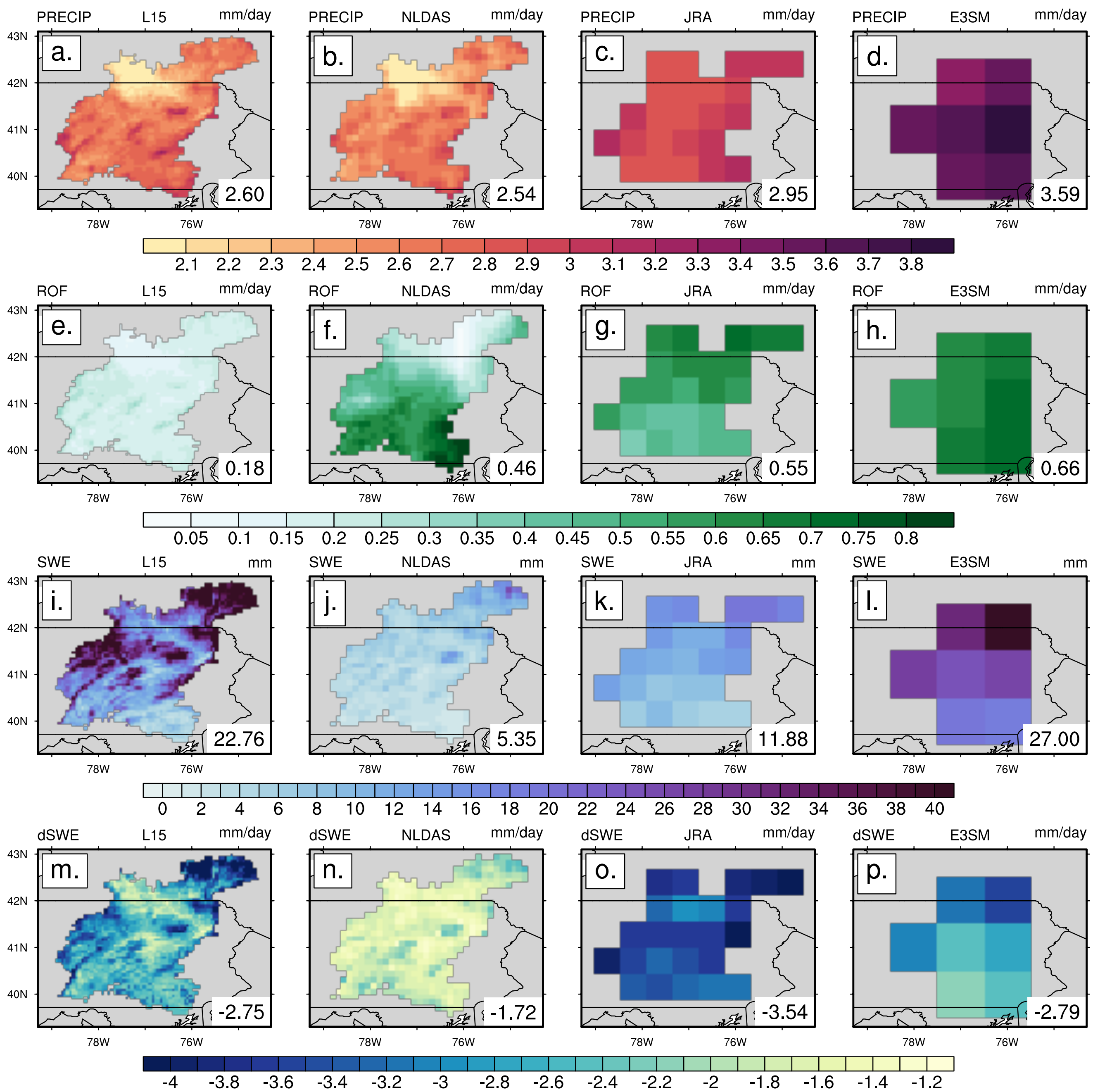

Figure 1. November to April (inclusive) mean climatologies of L15, NLDAS, JRA, and E3SM (left to right). From top to bottom are precipitation (PRECIP; mm/day), surface runoff (ROF; mm/day), snow water equivalent (SWE; mm), and daily change in snow water equivalent (dSWE; mm/day). dSWE only includes days with snow loss at a particular grid cell. Data is only included if >50% of a grid cell lies within the SRB bounds (black outline).

We focus our discussion on the relative threshold results in Fig. 3 for brevity, although note that the distributions of event statistics (Fig. 2b-d and Fig. 3b-d) are qualitatively similar between the two.

Figure 3a shows the total number of RoS events flagged over the climatological period (fourth column of Table 2), while Fig. 3b-d shows the frequency distribution for basin-averaged maximum rates of PRECIP, ROF, and dSWE at the event level. These are calculated by selecting the maximum daily value from the array of actual (i.e., unsmoothed) daily values during each RoS event.

The maximum PRECIP associated with flagged RoS events is similar between the four datasets, with E3SM tending to have slightly more extreme PRECIP occurring over the SRB (in agreement with Fig. 1). Much larger differences are seen in ROF and dSWE. In ROF, both the E3SM and JRA datasets produce larger event magnitudes of ROF, and subsequently have longer tails in their probability distributions, compared to the other two datasets. L15 produces RoS events with the least ROF (averaging approximately one-third of those found in either E3SM or JRA), with NLDAS in between the other three. The probability distribution functions of dSWE highlight an even more complex picture, with both NLDAS and L15 having narrower distributions with smaller magnitudes. Of note, JRA and E3SM have broader distributions (i.e., longer tails) but the JRA distribution is more skewed, with frequent low dSWE events whereas E3SM has a more uniform distribution of dSWE rates.

Summarizing, using either a relative or fixed threshold to identify RoS events, the fully-coupled ESMs (E3SM and JRA) produce more events than L15 and NLDAS. While daily PRECIP in each product differs somewhat, it should be noted that these differences are relatively small. Rather, the majority of the differences in RoS events flagged arise from lower magnitudes of ROF and dSWE in L15 and dSWE in NLDAS. It is worth noting that L15 and NLDAS produce fewer RoS events regardless of which thresholding technique is used, so not only are the distributions of relevant daily variables shifted relative to the other models, but their skewness is impacted as well (Figs. 2 and 3).

3.3 Single-year evaluation

To further explore RoS events at a more granular level we focus on results associated with the 1995 water year (WY95, October 1995 to September 1996). We chose this WY since this includes the January 1996 SRB flood that is often used for disaster planning purposes in the region (U.S. Army Corps of Engineers, 2001). The same visualizations for other WYs are included in the data download associated with this manuscript.

Figure 4 shows the SRB daily WY95 time series of basin-integrated PRECIP, ROF, and dSWE. Negative (positive) dSWE denotes snowmelt (accumulation). Vertical shading denotes RoS events that were flagged by the algorithm (using the RELATIVE thresholds). At the top of each panel is a stripe with four different shadings, representing days where streamflow exceeded the 75th, 90th, 95th, and 99th relative percentiles over the 1985-2005 period at the US Geological Survey (USGS) gage at Harrisburg, Pennsylvania (USGS #01570500), with deepest blues representing highly anomalous flow conditions.

The January 17-18, 1996 event is readily apparent for three of the four data products (NLDAS, JRA, and E3SM). Observed streamflow spiked during and immediately after the RoS event as shown by the dark blue stripes at the top of each time series, indicating a lag between when the RoS event algorithm identifies changes in PRECIP, dSWE, and ROF and the observed daily

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

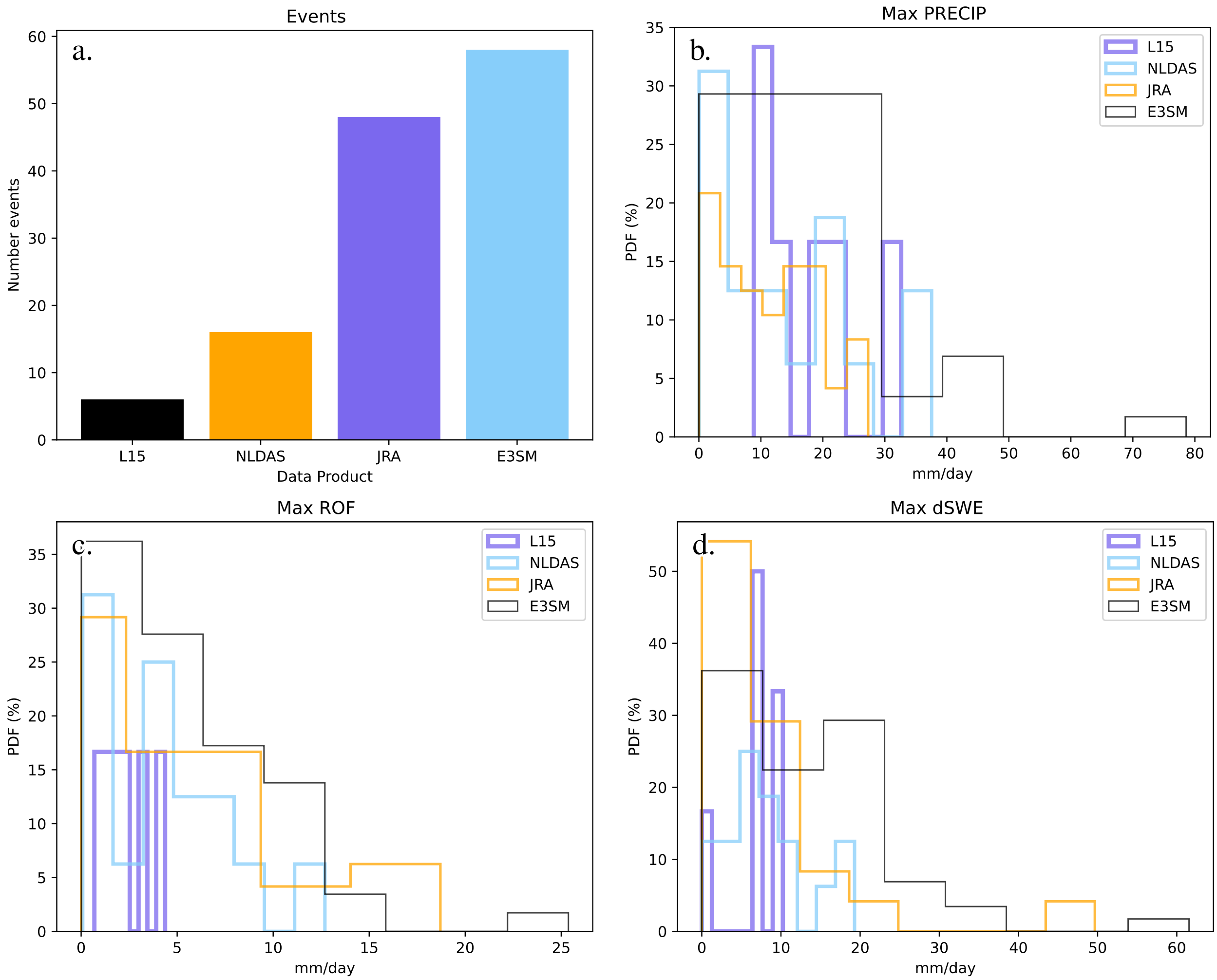

Figure 2. Histogram statistics for RoS events within the FIXED detection framework. The number of RoS events for each product from 1985-2005 is shown in the top left. The other three panels show frequency distributions of the daily rate (mm/day) of maximum precipitation (PRECIP), maximum runoff (ROF), and maximum change in SWE (dSWE, snowmelt positive).

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

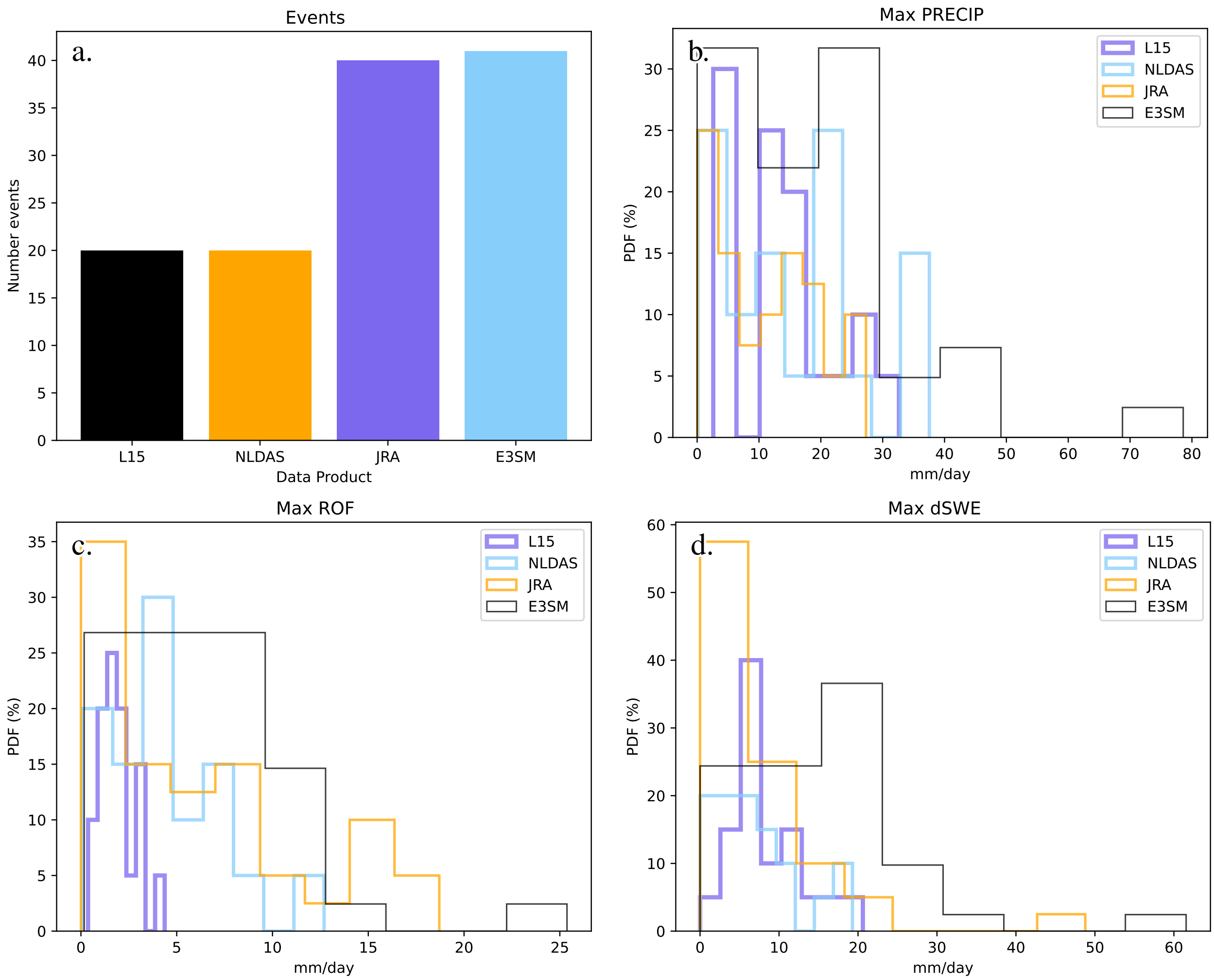

Figure 3. Same as Fig. 2 except for the RELATIVE detection thresholds.

streamflow gage observation at Harrisburg (namely, the 99th percentile streamflow exceedance). While this is operationally the largest RoS event observed over the period, there remain large discrepancies in the hydrometeorological conditions between the datasets. The dSWE signal is the largest in magnitude within JRA and E3SM compared to NLDAS. L15 shows a very minimal dSWE, to the point that t_{dSWE} is not exceeded in the detection algorithm. PRECIP is more similar across the three datasets, implying that the reduced ROF in L15 is likely due to minimal snowmelt.

All products also indicate a second, more moderate RoS event during the last third of February, although L15 and E3SM (JRA and E3SM) show larger dSWE (ROF) signals than the other two datasets. The USGS gage streamflow also exceeded the 95th percentile for this RoS event. More moderate flooding also appeared to occur in March and April although the detection algorithm only flagged such events in L15 and E3SM. We also emphasize that all RoS events flagged by the detection algo-

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

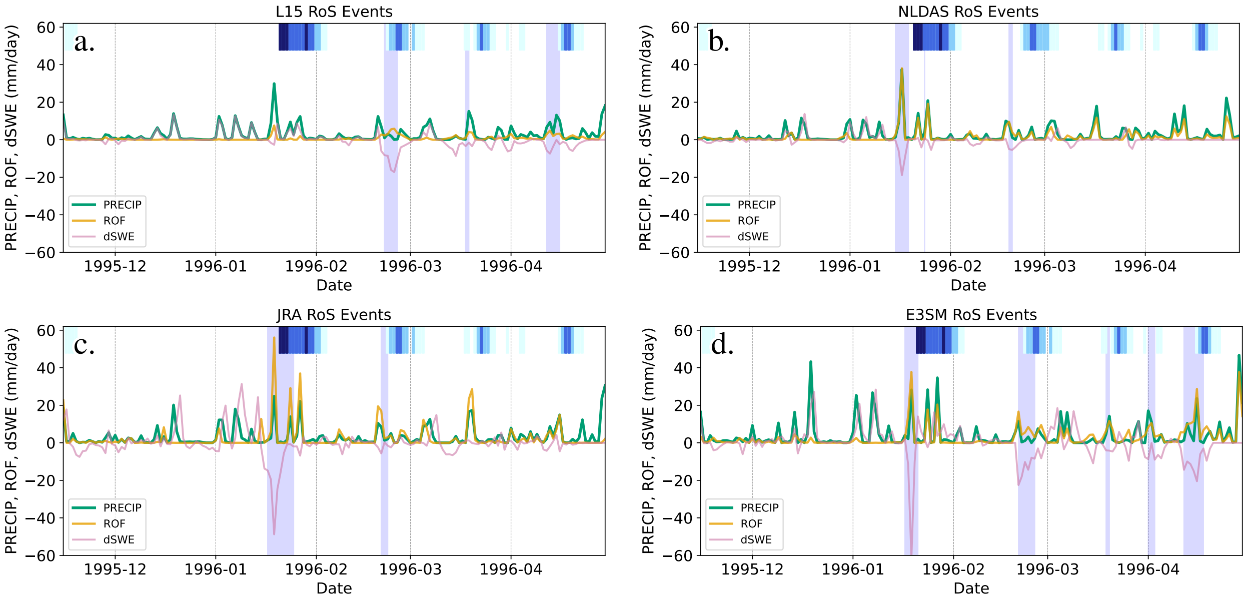

Figure 4. Time series of daily precipitation (PRECIP; green), surface runoff (ROF; orange), and SWE loss (dSWE; pink) during WY95 for each of the four products. Vertical purple shading denotes periods flagged as an RoS event. At the top of each time series are colored bars that denote extreme values (greater than the 75th percentile) of observed USGS streamflow at Harrisburg, Pennsylvania, with very dark hues representing 99th percentile values over the 1980-2005 period.

rithm resulted in well-above-average streamflow, highlighting the efficacy of the detection algorithm in capturing meaningful hydrometeorological extremes from basin-scale climate data.

Another way to visualize the WY95 RoS events is shown in Fig. 5. This plot shows the daily distribution of all three hydrometeorological quantities that make up a RoS event by plotting dSWE on the x-axis, PRECIP on the y-axis, and the size of the x-y marker depicts the magnitude of ROF. Markers are filled if the RoS event criteria are satisfied (RELATIVE) for that calendar day and are colored red (blue) if there is dSWE loss (gain). The variability of the PRECIP (y-axis) is the most similar between the two ESMs, although all products include at least one day of precipitation greater than 30 mm when averaged over the SRB. Larger differences across datasets are noted for the other two RoS event response variables. JRA and E3SM produce a wider spread in both RoS and non-RoS events along the y-axis, indicating that the two ESMs produce days with the largest magnitudes of dSWE. Further, L15 has markers that are small in size, indicating low daily ROF magnitudes when compared to the other three products. There is a substantial spread across the four datasets, even over comparable time series with well-defined and record-setting RoS events. This is likely due to daily differences in the timing of storms, the amount of water vapor transport and precipitation they produce, and how close surface air temperatures are to freezing conditions that in turn influence if a meteorological event leads to an increase in snow accumulation or results in an RoS event.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

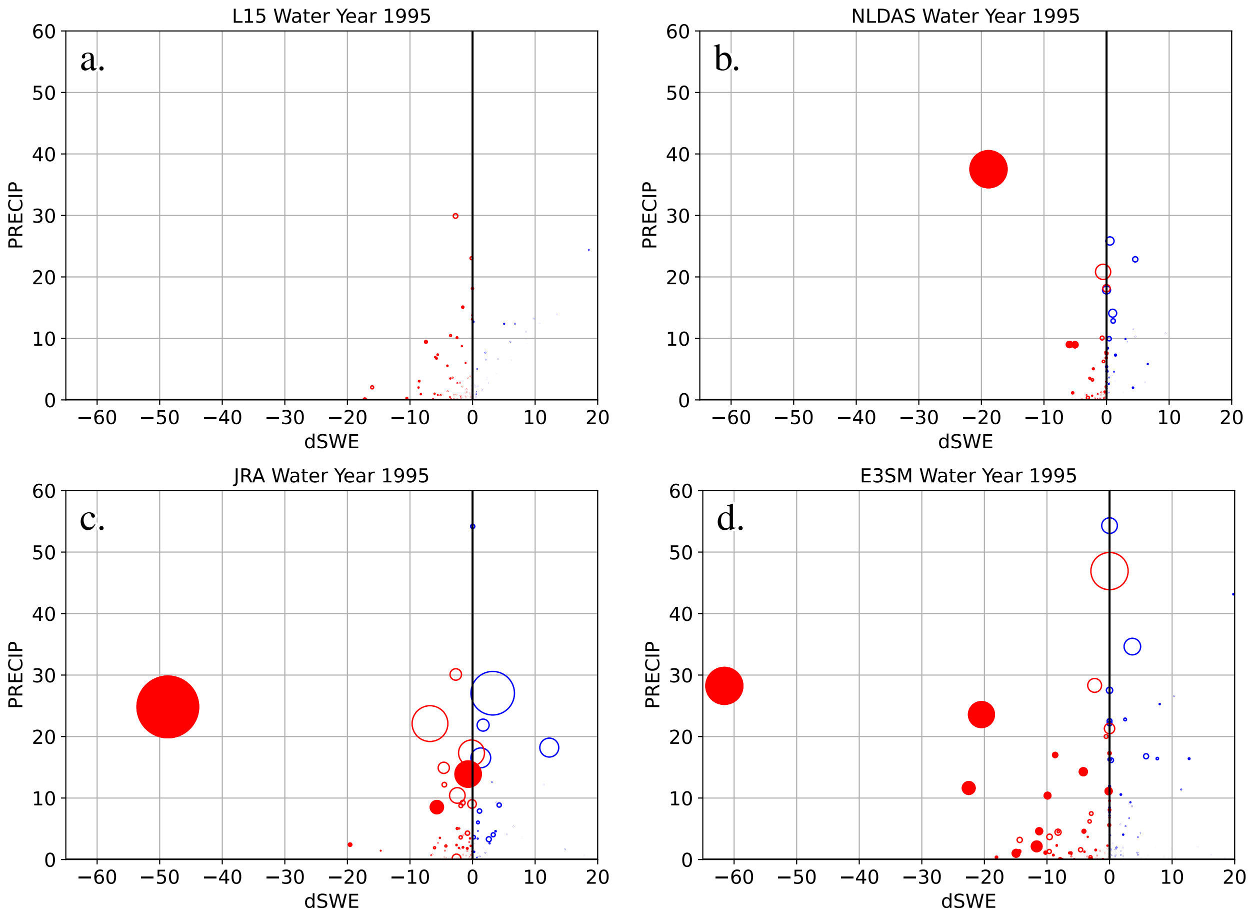

Figure 5. Bubble plots comparing precipitation (PRECIP; y-axis), SWE change (dSWE; x-axis), and surface runoff (size of bubble) for each of the four datasets during WY95. For each bubble, the values come from the same calendar day. Filled bubbles were flagged by the algorithm as an RoS event (purple shaded periods in Fig. 4).

3.4 Evaluation of a single event

One notable question given the above analysis (Fig. 4) is why the January 1996 event is captured differently across the products, particularly 'why is it 'unseen' by the L15 dataset?' Fig. 6 shows the SRB evolution of total on-surface SWE over WY95 for all datasets. All products show increasing SWE from mid-December onwards, albeit with varying accumulation rates and maximum depths. The datasets most differ in the SWE ablation during mid-January, with L15 (Fig. 6a) showing essentially no ablation of the existing snowpack, JRA (Fig. 6c) showing near total ablation, and NLDAS and E3SM (Fig. 6b,d) lying somewhere between the two extremes.

To untangle the potential source of this discrepancy, Fig. 7 shows the temperature (red line, right axis) at the grid cell nearest Harrisburg for the four products during the 1996 event. Also included is the temperature trace observed at Harrisburg (thin gold line) as recorded in National Oceanic and Atmospheric (NOAA) surface weather station archives along with a reference

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

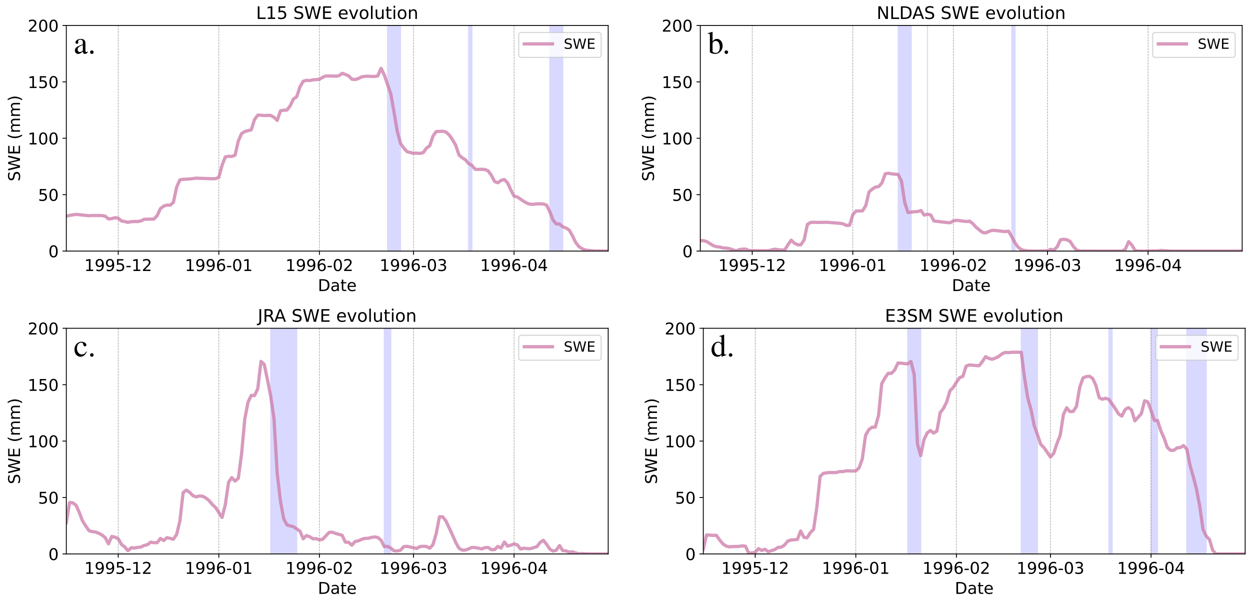

Figure 6. Daily SWE timeseries averaged over the SRB during WY95 for all four data products. Shaded in blue are the same events from Fig. 4.

freezing line (black). The surface relative humidity is indicated by the dashed purple line (right axis) and the product-specific associated precipitation rate (mm/hr) partitioned into rain (green) and snow (cyan) is denoted by bars (left axis).

None of the four products precisely match the observed temperature, although we acknowledge that this is, at least, partly due to the spatial resolution of the data. While the timing of the temperature increase associated with the warm air advection component of the event is well matched in NLDAS (Fig. 6b) and E3SM (Fig. 6d), both products have a slight cool bias. The overall temperature maximum is better matched by JRA (Fig. 6c), although the peak occurs approximately 6 hours earlier than in observations. An opposite shift is seen in L15 data, with the observed temperature peak leading L15's temperature peak by approximately 6 hours (Fig. 6a).

Across all four datasets, only L15 reports below-freezing temperatures between January 18th at 12Z and January 19th at 18Z. This occurs because L15 only leverages daily minimum and maximum observed temperatures to construct temperature fields. At each L15 grid cell, temperature minima and maxima are assumed to occur at sunrise and approximately late afternoon, respectively, with an interpolation spline used to provide information at intermediate hours (Bohn et al., 2013). These temperatures are always assumed to occur on the day of record (Livneh et al., 2015), which is relevant because the actual temperature minimum on January 19th occurred during the local afternoon after the passage of a cold front (Leathers et al., 1998). A similar, regular cycle is seen in the L15 relative humidity, with near-saturated surface conditions only occurring for approximately 3-hour windows each day. The other three data products show the increased moisture advection ahead of the precipitation maximum, with near-saturated surface air mass for large periods of the 72 hours preceding the initiation of the flood event. Therefore, for this event, the diurnal cycle of thermodynamic variables is more strongly influenced by atmospheric

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

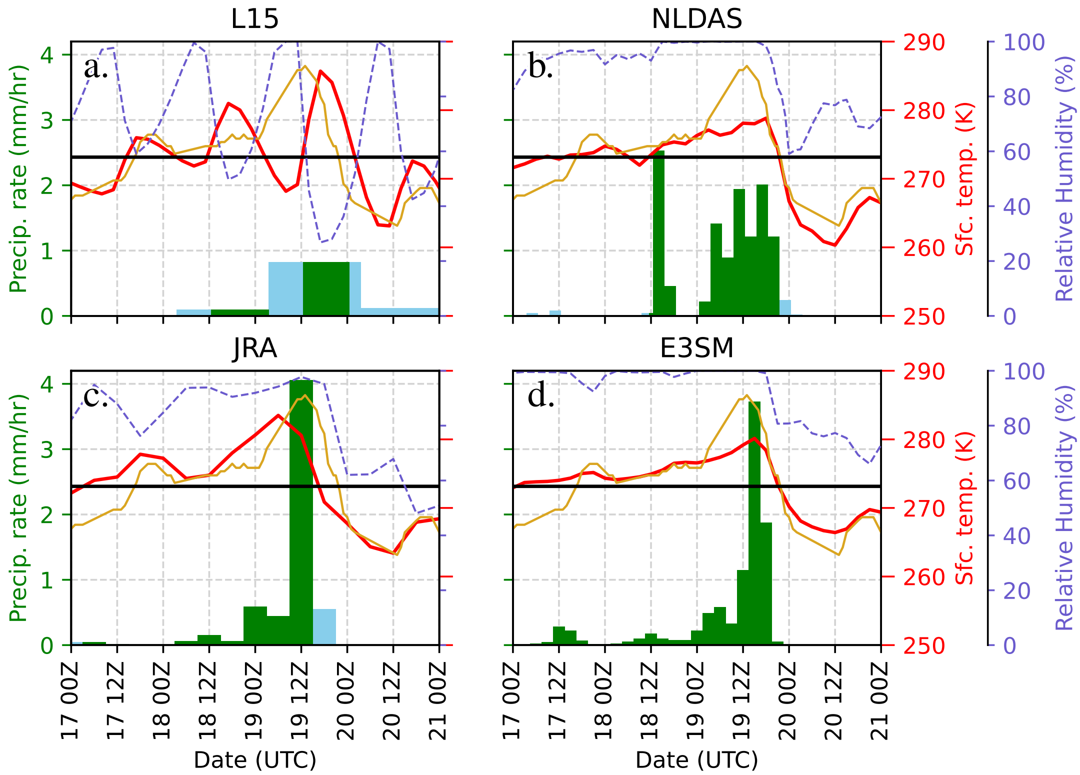

Figure 7. Evolution of reference height surface air temperature, relative humidity, and precipitation during the January 1996 SRB RoS flood event across data products. Data is taken from the grid cell nearest to Harrisburg, Pennsylvania. The left axis corresponds to the precipitation rate (shown in bars from the bottom) reported by the product. Green bars indicate rainfall (precipitation occurring when $T_s > 0°\mathrm{C}$) and cyan bars indicate snowfall (precipitation occurring when $T_s <= 0°\mathrm{C}$). The right axes show reference height surface air temperature as denoted by the solid red line, and relative humidity as denoted by the purple dashed line. A thick black horizontal line indicates 0°C air temperature (273.15 K). The thin gold line on each panel represents the observed temperature trace at Harrisburg over the period.

dynamics than solar insolation. We emphasize that this timing is actually quite important since factors such as temperature and surface humidity (and associated sensible and latent heat fluxes) are typically dominant drivers of snow ablation in high snowmelt events (Mazurkiewicz et al., 2008; Würzer et al., 2016; Harpold and Brooks, 2018) and need to line up concurrently to produce compound extremes like the one observed in 1996.

To understand how this could impact snowpack metamorphism, we also investigate precipitation during the event. We can assume precipitation type at the surface is determined by whether the reference height surface air temperature is above or below freezing, which is the manner in which the majority of land surface models partition precipitation (Harpold et al., 2017; Jennings et al., 2018; Siirila-Woodburn et al., 2021a). Since L15 contains below-freezing temperatures between 00Z and 12Z on January 19, it is the only product that partitions a fraction of precipitation into the frozen state during the peak of the event.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Further, since daily precipitation is evenly partitioned into sub-daily bins to force VIC, L15 registers more precipitation during this period than the other products (which typically peak after 12Z on January 19).

We argue this motivates the following interpretation. In L15, snowpack lost due to melt during the above-freezing temperatures during the 1996 event is, at least, partially offset by new snow on the ground that falls during the event due to below-freezing temperatures. These below-freezing periods and reduced relative humidity at the surface also mitigate any snowmelt in the land surface model that would be associated with enthalpy fluxes (either sensible or latent heat) into the existing snowpack. When combined, this explains the lack of a decline in L15 SWE and subsequent large spike in ROF in Fig. 4. All other datasets show a more prolonged period of above-freezing temperatures with high specific humidity during the window noted above, with nearly all precipitation during this period falling as rain versus snow. This induced varying degrees of snowmelt in E3SM, NLDAS, and JRA that are sufficient to trigger the RoS detection algorithm.

4 Conclusions

We interrogate four different gridded climate data products that include PRECIP, ROF, and SWE in order to quantify differences in the representation of coupled land-atmosphere processes that lead to flooding events in the historical record. In particular, we focus on RoS floods over the SRB and devise an algorithm for the automatic detection of RoS events that can be applied to any gridded dataset with relevant variables.

A detection algorithm flagging for times of collocated ROF and dSWE is generally successful at marking periods that will be followed by above-normal streamflow as measured by a gage in the basin. We find using fixed thresholds for RoS-relevant variables applied uniformly across multiple gridded datasets leads to large discrepancies in event frequency over the historical period, up to a factor of 10.

Normalizing these thresholds by each dataset's climatology (relative thresholds based on daily percentiles) improves agreement. However, there remains approximately a factor of two difference between event counts, implying that the underlying distributions are fundamentally different in both shape and magnitude across the data analyzed.

This underscores the complex assessment of such flood events – even products representing the historical record contain a large spread in hydrometeorological quantities provided to end users. The spatiotemporal dependencies of how meteorological data is generated for land surface and/or hydrological model forcing are critically important. While spatial resolution of data is important, particularly for snow processes in complex terrain (Henn et al., 2018; Siirila-Woodburn et al., 2021a), we show that the time resolution of data used to derive surface water conditions is just as critical for transient extremes, such as RoS events. This is because the timing of synoptic variations in reference height surface air temperature and humidity that dictate snowmelt and precipitation phase can occur on the order of hours at local scales and can have an outsized role in compound extreme event representation if those variations are threshold-dependent (e.g., storm rain-snow partitioning and alterations to the freezing line at the land surface).

Since RoS events in regions of ephemeral snow (e.g., SRB) can have rapid changes in surface forcing at hourly timescales (e.g., Leathers et al. (1998)), temporally coarse data applied to force LSMs may result in an underprediction of RoS frequency.

This may occur if time-interpolated data (and/or less frequent coupling) smooths and/or clips out short-term extrema and results in mismatched forcing or reduced day-to-day variability of multiple, co-dependent anomalies (e.g., temperature and precipitation extremes) required to accurately simulate historical, decision-relevant hydrometeorological extremes (e.g., 1996 SRB RoS event). We note that this effect can be independent of the model timestep itself. For example, the L15 data disaggregates precipitation data across sub-daily timesteps and uses a diurnal spline to reconstruct near-surface atmospheric conditions. While the LSM used is integrated at smaller timescales, the effective time resolution of the forcing data remains daily.

We also show that even though a dataset is provided at much coarser spatial resolutions than is desired (e.g., JRA and E3SM), model-derived datasets that are more frequently coupled and/or constrained at shorter timescales (e.g., reanalysis products and nudged ESMs) may produce more accurate land-atmosphere interactions and better representation of decision-relevant hydrometeorological extremes, particularly at basin (and larger) spatial scales. However, these products will likely suffer in the spatial representation of hydrometeorology in regions of high heterogeneity.

We want to emphasize that this work doesn't aim to suggest a 'superior' dataset writ large. For example, even though L15 does not capture the 1996 SRB flood as well as other products here, it has been an invaluable tool for studying climate extremes such as heat waves (Mazdiyasni and AghaKouchak, 2015), droughts (Pendergrass et al., 2020; Williams et al., 2020), and wildfires (Williams et al., 2019) (amongst others) in the US. Rather, we offer a few suggestions about leveraging hydrometeorological data for application purposes. It is recommended that gridded climate data developers consider the temporal resolution of land surface forcing if the representation of daily (or sub-daily) hydrometeorological extrema is desirable, particularly from a use-inspired or decision-relevant context (Jagannathan et al., 2021). While we do not downscale any datasets in this study, it is likely that using different meteorological data to force the same land surface and/or hydrological model will result in vastly different predicted streamflows, particularly for RoS events when variables are spatiotemporally co-dependent and would be sensitive to any post-processing adjustments (e.g., mean climatological correction). When possible, sub-daily information about atmospheric conditions should be included in meteorological forcing data.

From a stakeholder perspective, this is an important consideration when back-testing models, and, in particular, applying such models to evaluate tail risks (e.g., 1-in-100-year flood events and how they may change in a future climate). The results here show longer tails in the fully-coupled ESM-derived climate data, which may impact return rates of extreme events, even if calibrated or bias-corrected after the fact. Therefore, care must be taken when applying data requiring coupling between the atmosphere and land surface (and riverine) components, whether generated dynamically or statistically. While post-processing adjustments to the mean climatology may be desirable, these adjustments can alter decision-relevant hydrometeorological extremes that reside in the tail of the distribution.

Code and data availability. L15 data was downloaded from the NOAA Administration Physical Science Laboratory, available at https://psl.noaa.gov/data/gridded/data.livneh.html. NLDAS-VIC4.0.5 data was downloaded from the NOAA National Centers for Environmental Prediction, available at ftp://ldas.ncep.noaa.gov/nldas2/retrospective/vic4.0.5. JRA-55 data was downloaded from the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, available at https://doi.org/10.5065/

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

D6HH6H41. The version of E3SM run here was v2.0.0-alpha.2-1816-gf9cbe57a2 and is available at https://github.com/E3SM-Project/E3SM. ERA5 data used to nudge the E3SM solution was downloaded from the Copernicus Climate Data Store (CDS), available at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5. Station data for Harrisburg, PA was acquired from the Iowa Environmental Mesonet at https://mesonet.agron.iastate.edu/. The SRB is defined using a shapefile provided by the SRB Commission (https://www.srbc.net/portals/susquehanna-atlas/data-and-maps/subbasins/index.html). All processed data, figures, and scripts used to generate the results contained in this manuscript (along with README documentation) are archived at Zenodo and can be accessed at https://doi.org/10.5281/zenodo.10412332.

Author contributions. CZ devised the project and led the writing of the manuscript, with input from RM and AR. BA post-processed climate data and wrote the majority of the RoS detection algorithm. AR performed an initial cursory analysis and devised the bubble plot visualization.

Competing interests. The authors declare no competing interests.

Acknowledgements. This work is supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research program under Award DE-SC0016605 “A framework for improving analysis and modeling of Earth system and intersectoral dynamics at regional scales.” Data acquisition and E3SM simulations were completed at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award BER-ERCAP0020801. Event algorithm development, calibration, and analysis were performed on the Pennsylvania State University’s Institute for Computational and Data Sciences’ Roar supercomputer. The authors thank Dr. Andrew Jones for initial thoughts regarding the bubble plot diagnostics contained in this paper.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

References

AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., and Sadegh, M.: Climate Extremes and Compound Hazards in a Warming World, Annual Review of Earth and Planetary Sciences, 48, 519–548, https://doi.org/10.1146/annurev-earth-071719-055228, 2020.

Angélil, O., Stone, D., Wehner, M., Paciorek, C. J., Krishnan, H., and Collins, W.: An Independent Assessment of Anthropogenic Attribution Statements for Recent Extreme Temperature and Rainfall Events, Journal of Climate, 30, 5–16, https://doi.org/10.1175/JCLI-D-16-0077.1, 2017.

Ashley, S. T. and Ashley, W. S.: The storm morphology of deadly flooding events in the United States, International Journal of Climatology, 28, 493–503, https://doi.org/10.1002/joc.1554, 2008.

Beck, H. E., van Dijk, A. I., Roo, A. d., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Bohn, T. J., Livneh, B., Oyler, J. W., Running, S. W., Nijssen, B., and Lettenmaier, D. P.: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models, Agricultural and Forest Meteorology, 176, 38–49, https://doi.org/10.1016/j.agrformet.2013.03.003, 2013.

Brandt, W. T., Haleakala, K., Hatchett, B. J., and Pan, M.: A Review of the Hydrologic Response Mechanisms During Mountain Rain-on-Snow, Frontiers in Earth Science, 10, 791 760, https://doi.org/10.3389/feart.2022.791760, 2022.

Brown, R., Bartlett, P., MacKay, M., and Verseghy, D.: Evaluation of snow cover in CLASS for SnowMIP, Atmos.-Ocean, 44, 223–238, https://doi.org/10.3137/ao.440302, 2006.

Cohen, J., Ye, H., and Jones, J.: Trends and variability in rain-on-snow events, Geophysical Research Letters, 42, 7115–7122, https://doi.org/10.1002/2015GL065320, 2015.

Collins, M. J., Kirk, J. P., Pettit, J., DeGaetano, A. T., McCown, M. S., Peterson, T. C., Means, T. N., and Zhang, X.: Annual floods in New England (USA) and Atlantic Canada: synoptic climatology and generating mechanisms, Physical Geography, 35, 195–219, https://doi.org/10.1080/02723646.2014.888510, 2014.

Dougherty, E. and Rasmussen, K. L.: Climatology of flood-producing storms and their associated rainfall characteristics in the United States, Monthly Weather Review, 147, 3861–3877, https://doi.org/10.1175/MWR-D-19-0020.1, 2019.

Dougherty, E., Morrison, R., and Rasmussen, K.: High-resolution flood precipitation and streamflow relationships in two US river basins, Meteorol. Appl., 28, e1979, https://doi.org/10.1002/met.1979, 2021.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fasullo, J. T.: Evaluating simulated climate patterns from the CMIP archives using satellite and reanalysis datasets using the Climate Model Assessment Tool (CMATv1), Geoscientific Model Development, 13, 3627–3642, https://doi.org/10.5194/gmd-13-3627-2020, 2020.

Freudiger, D., Kohn, I., Stahl, K., and Weiler, M.: Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential, Hydrology and Earth System Sciences, 18, 2695–2709, https://doi.org/10.5194/hess-18-2695-2014, 2014.

George, S. St. and Mudelsee, M.: The weight of the flood-of-record in flood frequency analysis, Journal of Flood Risk Management, 12, e12 512, https://doi.org/10.1111/jfr3.12512, 2019.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Golaz, J.-C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., Bradley, A. M., Tang, Q., Maltrud, M. E., Forsyth, R. M., Zhang, C., Zhou, T., Zhang, K., Zender, C. S., Wu, M., Wang, H., Turner, A. K., Singh, B., Richter, J. H., Qin, Y., Petersen, M. R., Mametjanov, A., Ma, P.-L., Larson, V. E., Krishna, J., Keen, N. D., Jeffery, N., Hunke, E. C., Hannah, W. M., Guba, O., Griffin, B. M., Feng, Y., Engwirda, D., Di Vittorio, A. V., Dang, C., Conlon, L. M., Chen, C.-C.-J., Brunke, M. A., Bisht, G., Benedict, J. J., Asay-Davis, X. S., Zhang, Y., Zhang, M., Zeng, X., Xie, S., Wolfram, P. J., Vo, T., Veneziani, M., Tesfa, T. K., Sreepathi, S., Salinger, A. G., Eyre, J. E. J. R., Prather, M. J., Mahajan, S., Li, Q., Jones, P. W., Jacob, R. L., Huebler, G. W., Huang, X., Hillman, B. R., Harrop, B. E., Foucar, J. G., Fang, Y., Comeau, D. S., Caldwell, P. M., Bartoletti, T., Balaguru, K., Taylor, M. A., McCoy, R. B., Leung, L. R., and Bader, D. C.: The DOE E3SM Model Version 2: Overview of the Physical Model and Initial Model Evaluation, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 156, https://doi.org/10.1029/2022MS003156, 2022.

Grote, T.: A synoptic climatology of rain-on-snow flooding in Mid-Atlantic region using NCEP/NCAR Re-Analysis, Physical Geography, 42, 452–471, https://doi.org/10.1080/02723646.2020.1838119, 2021.

Guan, B., Waliser, D. E., Ralph, F. M., Fetzer, E. J., and Neiman, P. J.: Hydrometeorological characteristics of rain-on-snow events associated with atmospheric rivers, Geophysical Research Letters, 43, 2964–2973, https://doi.org/10.1002/2016GL067978, 2016.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., et al.: Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe, Journal of Hydrometeorology, 13, 604–620, https://doi.org/10.1175/JHM-D-11-083.1, 2012.

Gutmann, E. D., Rasmussen, R. M., Liu, C., Ikeda, K., Gochis, D. J., Clark, M. P., Dudhia, J., and Thompson, G.: A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation over Complex Terrain, Journal of Climate, 25, 262–281, https://doi.org/10.1175/2011JCLI4109.1, 2012.

Haleakala, K., Brandt, W. T., Hatchett, B. J., Li, D., Lettenmaier, D. P., and Gebremichael, M.: Watershed memory amplified the Oroville rain-on-snow flood of February 2017, PNAS Nexus, 2, pgac295, https://doi.org/10.1093/pnasnexus/pgac295, 2023.

Harpold, A. A. and Brooks, P. D.: Humidity determines snowpack ablation under a warming climate, Proceedings of the National Academy of Sciences, 115, 1215–1220, https://doi.org/10.1073/pnas.1716789115, 2018.

Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., and Steele, C. M.: Rain or snow: hydrologic processes, observations, prediction, and research needs, Hydrology and Earth System Sciences, 21, 1–22, https://doi.org/10.5194/hess-21-1-2017, 2017.

Hartley, S. and Keables, M. J.: Synoptic associations of winter climate and snowfall variability in New England, USA, 1950-1992, International Journal of Climatology, 18, 281–298, https://doi.org/10.1002/(SICI)1097-0088(19980315)18:3<281::AID-JOC245>3.0.CO;2-F, 1998.

Hatchett, B. J.: Seasonal and Ephemeral Snowpacks of the Conterminous United States, Hydrology, 8, https://doi.org/10.3390/hydrology8010032, 2021.

Heggli, A., Hatchett, B., Schwartz, A., Bardsley, T., and Hand, E.: Toward snowpack runoff decision support, iScience, 25, 104 240, https://doi.org/10.1016/j.isci.2022.104240, 2022.

Henn, B., Newman, A. J., Livneh, B., Daly, C., and Lundquist, J. D.: An assessment of differences in gridded precipitation datasets in complex terrain, Journal of Hydrology, 556, 1205–1219, https://doi.org/10.1016/j.jhydrol.2017.03.008, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J.,

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hill, C., DeLuca, C., Suarez, M., Da Silva, A., et al.: The architecture of the Earth System Modeling Framework, Computing in Science & Engineering, 6, 18–28, 2004.

Holton, J. R.: An Introduction to Dynamic Meteorology, Elsevier Acadamic Press, fourth edn., ISBN 0123540151, 535 pp., 2004.

Hotovy, O., Nedelcev, O., and Jenicek, M.: Changes in rain-on-snow events in mountain catchments in the rain-snow transition zone, Hydrological Sciences Journal, https://doi.org/10.1080/02626667.2023.2177544, 2023.

Jagannathan, K., Jones, A. D., and Ray, I.: The Making of a Metric: Co-Producing Decision-Relevant Climate Science, Bulletin of the American Meteorological Society, 102, E1579 – E1590, https://doi.org/10.1175/BAMS-D-19-0296.1, 2021.

Jennings, K. S., Winchell, T. S., Livneh, B., and Molotch, N. P.: Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere, Nature Communications, 9, 1–9, https://doi.org/10.1038/s41467-018-03629-7, 2018.

Jeong, D. I. and Sushama, L.: Rain-on-snow events over North America based on two Canadian regional climate models, Climate Dynamics, 50, 303–316, https://doi.org/10.1007/s00382-017-3609-x, 2018.

Knowles, N., Dettinger, M. D., and Cayan, D. R.: Trends in snowfall versus rainfall in the western United States, Journal of Climate, 19, 4545–4559, 2006.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., et al.: The JRA-55 reanalysis: General specifications and basic characteristics, Journal of the Meteorological Society of Japan. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Leathers, D. J., Kluck, D. R., and Kroczynski, S.: The severe flooding event of January 1996 across north-central Pennsylvania, Bulletin of the American Meteorological Society, 79, 785–798, https://doi.org/10.1175/1520-0477(1998)079<0785:TSFEOJ>2.0.CO;2, 1998.

Leathers, D. J., Malin, M. L., Kluver, D. B., Henderson, G. R., and Bogart, T. A.: Hydroclimatic variability across the Susquehanna River Basin, USA, since the 17th century, International Journal of Climatology, 28, 1615–1626, https://doi.org/10.1002/joc.1668, 2008.

Li, D., Lettenmaier, D. P., Margulis, S. A., and Andreadis, K.: The Role of Rain-on-Snow in Flooding Over the Conterminous United States, Water Resources Research, 55, 8492–8513, https://doi.org/10.1029/2019WR024950, 2019.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research: Atmospheres, 99, 14 415–14 428, https://doi.org/10.1029/94JD00483, 1994.

Livneh, B., Deems, J. S., Schneider, D., Barsugli, J. J., and Molotch, N. P.: Filling in the gaps: Inferring spatially distributed precipitation from gauge observations over complex terrain, Water Resources Research, 50, 8589–8610, https://doi.org/10.1002/2014WR015442, 2014.

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D. R., and Brekke, L.: A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013, Scientific Data, 2, 1–12, https://doi.org/10.1038/sdata.2015.42, 2015.

López-Moreno, J. I., Pomeroy, J. W., Morán-Tejeda, E., Revuelto, J., Navarro-Serrano, F. M., Vidaller, I., and Alonso-González, E.: Changes in the frequency of global high mountain rain-on-snow events due to climate warming, Environmental Research Letters, 16, 094 021, https://doi.org/10.1088/1748-9326/ac0dde, 2021.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Lundquist, J. D., Hughes, M., Henn, B., Gutmann, E. D., Livneh, B., Dozier, J., and Neiman, P.: High-Elevation Precipitation Patterns: Using Snow Measurements to Assess Daily Gridded Datasets across the Sierra Nevada, California, Journal of Hydrometeorology, 16, 1773–1792, https://doi.org/10.1175/JHM-D-15-0019.1, 2015.

Maina, F. Z. and Kumar, S. V.: Diverging Trends in Rain-On-Snow Over High Mountain Asia, Earth's Future, 11, e2022EF003 009, https://doi.org/10.1029/2022EF003009, 2023.

Mazdiyasni, O. and AghaKouchak, A.: Substantial increase in concurrent droughts and heatwaves in the United States, Proceedings of the National Academy of Sciences, 112, 11 484–11 489, https://doi.org/10.1073/pnas.1422945112, 2015.

Mazurkiewicz, A. B., Callery, D. G., and McDonnell, J. J.: Assessing the controls of the snow energy balance and water available for runoff in a rain-on-snow environment, Journal of Hydrology, 354, 1–14, https://doi.org/10.1016/j.jhydrol.2007.12.027, 2008.

McCabe, G. J. and Clark, M. P.: Trends and variability in snowmelt runoff in the western United States, Journal of Hydrometeorology, 6, 476–482, 2005.

McCabe, G. J., Clark, M. P., and Hay, L. E.: Rain-on-snow events in the western United States, Bulletin of the American Meteorological Society, 88, 319–328, https://doi.org/10.1175/BAMS-88-3-319, 2007.

McCrary, R. R., McGinnis, S., and Mearns, L. O.: Evaluation of Snow Water Equivalent in NARCCAP Simulations, Including Measures of Observational Uncertainty, Journal of Hydrometeorology, 18, 2425–2452, https://doi.org/10.1175/JHM-D-16-0264.1, 2017.

McCrary, R. R., Mearns, L. O., Hughes, M., Biner, S., and Bukovsky, M. S.: Projections of North American snow from NA-CORDEX and their uncertainties, with a focus on model resolution, Clim. Change, 170, 1–25, https://doi.org/10.1007/s10584-021-03294-8, 2022.

Moore, R. and Owens, I.: Controls on advective snowmelt in a maritime alpine basin, Journal of Applied Meteorology and Climatology, 23, 135–142, https://doi.org/10.1175/1520-0450(1984)023<0135:COASIA>2.0.CO;2, 1984.

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., and Rasmussen, R.: Projected increases and shifts in rain-on-snow flood risk over western North America, Nature Climate Change, 8, 808–812, https://doi.org/10.1038/s41558-018-0236-4, 2018.

Niu, G.-Y. and Yang, Z.-L.: Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, Journal of Hydrometeorology, 7, 937–952, 2006.

Parker, W. S.: Reanalyses and Observations: What's the Difference?, Bulletin of the American Meteorological Society, 97, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1, 2016.

Parker, W. S.: Model Evaluation: An Adequacy-for-Purpose View, Philosophy of Science, 87, 457–477, https://doi.org/10.1086/708691, 2020.

Pendergrass, A. G., Meehl, G. A., Pulwarty, R., Hobbins, M., Hoell, A., AghaKouchak, A., Bonfils, C. J. W., Gallant, A. J. E., Hoerling, M., Hoffmann, D., Kaatz, L., Lehner, F., Llewellyn, D., Mote, P., Neale, R. B., Overpeck, J. T., Sheffield, A., Stahl, K., Svoboda, M., Wheeler, M. C., Wood, A. W., and Woodhouse, C. A.: Flash droughts present a new challenge for subseasonal-to-seasonal prediction, Nature Climate Change, 10, 191–199, https://doi.org/10.1038/s41558-020-0709-0, 2020.

Perry, C. A.: Significant floods in the United States during the 20th century: USGS measures a century of floods, US Department of the Interior, US Geological Survey, 2000.

Pettett, A. and Zarzycki, C. M.: The 1996 Mid-Atlantic Winter Flood: Exploring Climate Risk through a Storyline Approach, Journal of Hydrometeorology, 24, 2259–2280, https://doi.org/10.1175/JHM-D-22-0146.1, 2023.

Poschlod, B., Zscheischler, J., Sillmann, J., Wood, R. R., and Ludwig, R.: Climate change effects on hydrometeorological compound events over southern Norway, Weather and Climate Extremes, 28, 100 253, https://doi.org/10.1016/j.wace.2020.100253, 2020.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

Pradhanang, S. M., Frei, A., Zion, M., Schneiderman, E. M., Steenhuis, T. S., and Pierson, D.: Rain-on-snow runoff events in New York, Hydrological Processes, 27, 3035–3049, https://doi.org/10.1002/hyp.9864, 2013.

Rauscher, S. A., Pal, J. S., Diffenbaugh, N. S., and Benedetti, M. M.: Future changes in snowmelt-driven runoff timing over the western US, Geophysical Research Letters, 35, 2008.

Rhoades, A. M., Jones, A. D., and Ullrich, P. A.: Assessing Mountains as Natural Reservoirs With a Multimetric Framework, Earth's Future, 6, 1221–1241, https://doi.org/10.1002/2017EF000789, 2018.

Sharma, S., Gomez, M., Keller, K., Nicholas, R. E., and Mejia, A.: Regional Flood Risk Projections under Climate Change, Journal of Hydrometeorology, 22, 2259–2274, https://doi.org/10.1175/JHM-D-20-0238.1, 2021.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., et al.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic change, 151, 555–571, https://doi.org/10.1007/s10584-018-2317-9, 2018.

Siirila-Woodburn, E., Rhoades, A. M., Hatchett, B. J., Huning, L., Szinai, J., Tague, C., Nico, P. S., Feldman, D., Jones, A. D., Collins, W. D., and Kaatz, L.: A low-to-no snow future and its impacts on water resources in the western United States, Nature Reviews Earth and Environment, 2, 800–819, https://doi.org/10.1038/s43017-021-00219-y, 2021a.

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., Collins, W. D., and Kaatz, L.: A low-to-no snow future and its impacts on water resources in the western United States, Nature Reviews Earth & Environment, 2, 800–819, https://doi.org/10.1038/s43017-021-00219-y, 2021b.

Singh, P., Spitzbart, G., Hübl, H., and Weinmeister, H.: Hydrological response of snowpack under rain-on-snow events: a field study, Journal of Hydrology, 202, 1–20, https://doi.org/10.1016/S0022-1694(97)00004-8, 1997.

Sood, A. and Smakhtin, V.: Global hydrological models: a review, Hydrological Sciences Journal, 60, 549–565, 2015.

Sterle, K., Hatchett, B. J., Singletary, L., and Pohll, G.: Hydroclimate Variability in Snow-Fed River Systems: Local Water Managers' Perspectives on Adapting to the New Normal, Bulletin of the American Meteorological Society, 100, 1031–1048, https://doi.org/10.1175/BAMS-D-18-0031.1, 2019.

Sun, J., Zhang, K., Wan, H., Ma, P.-L., Tang, Q., and Zhang, S.: Impact of Nudging Strategy on the Climate Representativeness and Hindcast Skill of Constrained EAMv1 Simulations, Journal of Advances in Modeling Earth Systems, 11, 3911–3933, https://doi.org/10.1029/2019MS001831, 2019.

Surfleet, C. G. and Tullos, D.: Variability in effect of climate change on rain-on-snow peak flow events in a temperate climate, Journal of Hydrology, 479, 24–34, https://doi.org/10.1016/j.jhydrol.2012.11.021, 2013.

Suriano, Z. J.: North American rain-on-snow ablation climatology, Climate Research, 87, 133–145, https://doi.org/10.3354/cr01687, 2022.

Suriano, Z. J., Henderson, G. R., and Leathers, D. J.: Discharge responses associated with rapid snow cover ablation events in the Susquehanna and Wabash River basins, Physical Geography, 41, 70–82, https://doi.org/10.1080/02723646.2019.1674558, 2020.

Suriano, Z. J., Henderson, G. R., Arthur, J., Harper, K., and Leathers, D. J.: Atmospheric Drivers Associated with Extreme Snow Ablation and Discharge Events in the Susquehanna River Basin: A Climatology, Journal of Applied Meteorology and Climatology, 62, 1497–1510, https://doi.org/10.1175/JAMC-D-23-0042.1, 2023.

Toomey, M., Cantwell, M., Colman, S., Cronin, T., Donnelly, J., Giosan, L., Heil, C., Korty, R., Marot, M., and Willard, D.: The Mighty Susquehanna – Extreme Floods in Eastern North America During the Past Two Millennia, Geophysical Research Letters, 46, 3398–3407, https://doi.org/10.1029/2018GL080890, 2019.

https://doi.org/10.5194/egusphere-2023-3094
Preprint. Discussion started: 14 March 2024

U.S. Army Corps of Engineers: Non-Structural Flood Damage Reduction Within the Corps of Engineers: What Districts Are Doing, Tech. Rep. ADA629409, https://apps.dtic.mil/sti/tr/pdf/ADA629409.pdf, 2001.

Villarini, G. and Smith, J. A.: Flood peak distributions for the eastern United States, Water Resources Research, 46, https://doi.org/10.1029/2009WR008395, 2010.

Wachowicz, L. J., Mote, T. L., and Henderson, G. R.: A rain on snow climatology and temporal analysis for the eastern United States, Physical Geography, 41, 54–69, https://doi.org/10.1080/02723646.2019.1629796, 2020.

Wayand, N. E., Lundquist, J. D., and Clark, M. P.: Modeling the influence of hypsometry, vegetation, and storm energy on snowmelt contributions to basins during rain-on-snow floods, Water Resources Research, 51, 8551–8569, https://doi.org/10.1002/2014WR016576, 2015.

Williams, A. P., Abatzoglou, J. T., Gershunov, A., Guzman-Morales, J., Bishop, D. A., Balch, J. K., and Lettenmaier, D. P.: Observed Impacts of Anthropogenic Climate Change on Wildfire in California, Earth's Future, 7, 892–910, https://doi.org/10.1029/2019EF001210, 2019.

Williams, A. P., Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles, K., Baek, S. H., Badger, A. M., and Livneh, B.: Large contribution from anthropogenic warming to an emerging North American megadrought, Science, 368, 314–318, https://doi.org/10.1126/science.aaz9600, 2020.

Würzer, S., Jonas, T., Wever, N., and Lehning, M.: Influence of initial snowpack properties on runoff formation during rain-on-snow events, Journal of Hydrometeorology, 17, 1801–1815, https://doi.org/10.1175/JHM-D-15-0181.1, 2016.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., et al.: 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, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2011JD016048, 2012.

Ye, H., Yang, D., and Robinson, D.: Winter rain on snow and its association with air temperature in northern Eurasia, Hydrological Processes, 22, 2728–2736, https://doi.org/10.1002/hyp.7094, 2008.

Yu, G., Wright, D. B., and Davenport, F. V.: Diverse Physical Processes Drive Upper-Tail Flood Quantiles in the US Mountain West, Geophysical Research Letters, 49, e2022GL098 855, https://doi.org/10.1029/2022GL098855, 2022.

Zarzycki, C. M. and Jablonowski, C.: Experimental Tropical Cyclone Forecasts Using a Variable-Resolution Global Model, Monthly Weather Review, 143, 4012–4037, https://doi.org/10.1175/MWR-D-15-0159.1, 2015.

Zhang, S., Zhang, K., Wan, H., and Sun, J.: Further improvement and evaluation of nudging in the E3SM Atmosphere Model version 1 (EAMv1): simulations of the mean climate, weather events, and anthropogenic aerosol effects, Geoscientific Model Development, 15, 6787–6816, https://doi.org/10.5194/gmd-15-6787-2022, 2022.

Zheng, H., Yang, Z.-L., Lin, P., Wei, J., Wu, W.-Y., Li, L., Zhao, L., and Wang, S.: On the Sensitivity of the Precipitation Partitioning Into Evapotranspiration and Runoff in Land Surface Parameterizations, Water Resour. Res., 55, 95–111, https://doi.org/10.1029/2017WR022236, 2019.