Mass-Conserving Downscaling of Climate Model Precipitation Over Mountainous Terrain for Water Resource Applications

A. Rugg[1] ⓘ, E. D. Gutmann[1] ⓘ, R. R. McCrary[1], F. Lehner[1,2,3] ⓘ, A. J. Newman[1] ⓘ, J. H. Richter[1] ⓘ, M. R. Tye[1,4] ⓘ, and A. W. Wood[1] ⓘ

[1]National Center for Atmospheric Research, Boulder, CO, USA, [2]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, USA, [3]Polar Bears International, Bozeman, MT, USA, [4]Whiting School of Civil Engineering, Johns Hopkins University, Baltimore, MD, USA

Abstract A mass-conserving method to downscale precipitation from global climate models (GCMs) using sub-grid-scale topography and modeled 700-mb wind direction is presented. Runoff is simulated using a stand-alone hydrological model, with this and several other methods as inputs, and compared to runoff simulated using historical observations over the western contiguous United States. Results suggest the mitigation of grid-scale biases is more critical than downscaling for some regions with large wet biases (e.g., the Great Basin and Upper Colorado). In other regions (e.g., the Pacific Northwest) the new method produces more realistic sub-grid-scale variability in runoff compared to unadjusted GCM output and a simpler downscaling method. The presented method also brings the runoff centroid timing closer to that simulated with observations for all subregions examined.

Plain Language Summary Due to limitations in computing power which necessitates coarse spatial resolution, climate models do not include many details on mountains and their impact on precipitation. For this reason, it is difficult to estimate the impact of climate change on the availability of water for human consumption in places like the western United States, where mountain snowpack is an important source of water. This paper presents a way to adjust precipitation estimates from climate models by using some simple statistics about nearby mountains and valleys. The adjusted precipitation is then used in a hydrologic model to calculate the runoff that is simulated with different precipitation inputs. Results show that in most cases, the precipitation adjustment improves estimates of the resultant runoff relative to simulations with observed precipitation. The main exceptions are in very dry areas where the climate model produces far too much precipitation. With further work, the proposed adjustment may be integrated into the climate model itself to make it easier for those managing water resources (e.g., controlling reservoir levels) to use the model output to plan and adapt to climate change.

1. Introduction

A 2°C increase in global mean temperature is expected to cause important changes in the water cycle for over half of the world's population (Sedláček & Knutti, 2014). Global climate models (GCMs) can be used to inform long-term water resource planning, but their accuracy is hindered by their coarse resolution (e.g., Ekström et al., 2018; Null et al., 2010) and biases in their mean state and sensitivities (Lehner et al., 2019). This is especially true for mountainous regions and downstream lowlands which depend on the seasonal melting of snowpack. While the demand for water in these areas is expected to increase in coming decades, the amount of water stored in snowpack is expected to decrease due to climate change (Viviroli et al., 2006, 2011, 2020).

The severity of these climate-change-driven water cycle changes are highly uncertain since mountainous terrain is not well resolved by the roughly 100-km grid spacing of GCMs, especially since minimum skillful scale of a GCM is greater than the actual grid box size (e.g., Takayabu et al., 2016). The relevant hydrometeorological processes are better captured by higher-resolution simulations, but such simulations are too computationally expensive to run globally for centuries (e.g., Chan et al., 2012; Prein et al., 2013, 2015). Many statistical methods for downscaling precipitation from GCMs exist that are computationally efficient (e.g., Gutmann et al., 2022; Maurer & Hidalgo, 2008; Pierce et al., 2014; Wood et al., 2004), but these methods are often tuned to small regions or individual grid-cells and are not necessarily applicable to data-sparse regions. Furthermore, these methods are not mass-conserving, which means they cannot be coupled with the GCM to provide feedback to

Resources: E. D. Gutmann, R. R. McCrary
Software: R. R. McCrary
Supervision: E. D. Gutmann
Validation: A. Rugg
Visualization: A. Rugg
Writing – original draft: A. Rugg
Writing – review & editing: A. Rugg, E. D. Gutmann, R. R. McCrary, F. Lehner, A. J. Newman, J. H. Richter, M. R. Tye, A. W. Wood

the model. In this paper we develop a generalizable downscaling technique that is computationally inexpensive, mass-conserving, and may be coupled with sub-grid-scale land components of GCMs.

The enhancement and reduction of precipitation on windward and leeward slopes, respectively, has been described quantitatively since the 1970s (e.g., Collier, 1975; Hobbs et al., 1973; Smith & Barstad, 2004). Our method captures these orographic effects by using the GCM-modeled 700-mb wind direction and upwind topography information. Wind direction has previously been used to downscale precipitation over mountainous terrain on a regional scale (e.g., Salathé Jr, 2003; Schipper et al., 2011) and wind speed on a continental scale (Ghan & Shippert, 2006; Tesfa et al., 2020). We expand upon previous work to disaggregate daily minimum and maximum temperature (TMin and TMax) and daily precipitation (P) using sub-grid-scale topographical features. We focus on the western contiguous United States because of the complex terrain, large variety of climates across the region, and availability of historical TMin, TMax, and P temperature observations (Livneh et al., 2013). We apply our method to TMin, TMax, and P from the Community Earth System Model v.2 (CESM2, Danabasoglu et al., 2020) and estimate the impact on runoff using the Variable Infiltration Capacity hydrological model (VIC, Liang et al., 1994).

Using VIC we compare simulated total annual runoff and the seasonal timing of that runoff for subregions of the western United States. The total annual runoff is necessary for balancing the overall water budget of a region, and the timing of that runoff has a large impact on flood risk, reservoir management, and hydropower availability. For example, snowmelt is expected to produce earlier peak runoff in much of the western United States, meaning more water will need to be stored in reservoirs to remain useable in summer and autumn when agricultural demands are highest (e.g., Barnett et al., 2005; Li et al., 2017).

The focus of this paper is a downscaling method which considers the elevation both at the grid point of interest and upwind of that gridpoint. To quantify the value of this "Full" adjustment on modeled runoff we use VIC to simulate runoff using TMin, TMax, and P from:

- Observations ("Livneh");
- CESM2 with no adjustments ("Raw");
- CESM2 with a mass-conserving adjustment based on elevation at a given grid point ("Elev");
- CESM2 with a mass-conserving adjustment based on elevation and wind ("Full"); and
- CESM2 with a non-mass-conserving downscaling and seasonal bias-correction technique ("Full-BC").

These runs allow us to quantify the impact of the gridpoint elevation adjustment alone, the upwind topography terms included in the Full adjustment, and the Full adjustment if seasonal grid-scale biases in the CESM2 are corrected. Comparisons between these simulations reveal the relative importance of large-scale bias correction and spatial downscaling to the utility of GCM output.

2. Data

2.1. Historical Observations

The Livneh et al. (2013) data set is a gridded product (1/16th degree) created from the interpolated daily observations with adjustments for climatological and topographic effects. The data set is hereafter referred to as the Livneh data set.

2.2. Global Climate Model Output

The downscaling and bias correction techniques are applied to output from the CESM2 Large Ensemble Project (Danabasoglu et al., 2020; Rodgers et al., 2021). Ensemble members differ only in the initialization year (between 1001 and 1231), and small perturbations (order 10^{-14} K) made to the initial atmospheric potential temperature field (Rodgers et al., 2021), which is subject to historical forcing (1850–2014). Historical output from ensemble members 1–10 were compared to Livneh to derive coefficients for the adjustments described in Section 3. Those adjustments were then applied to TMin, TMax, and P historical output of ensemble member 11 for an independent sample. All results shown are for ensemble member 11 alone. While the adjustments are applied to daily CESM2 output, the evaluation of them (Section 4) relies on the comparison of statistics averaged over 30 years since GCMs are not designed to reproduce the chronology of historical events on shorter timescales.
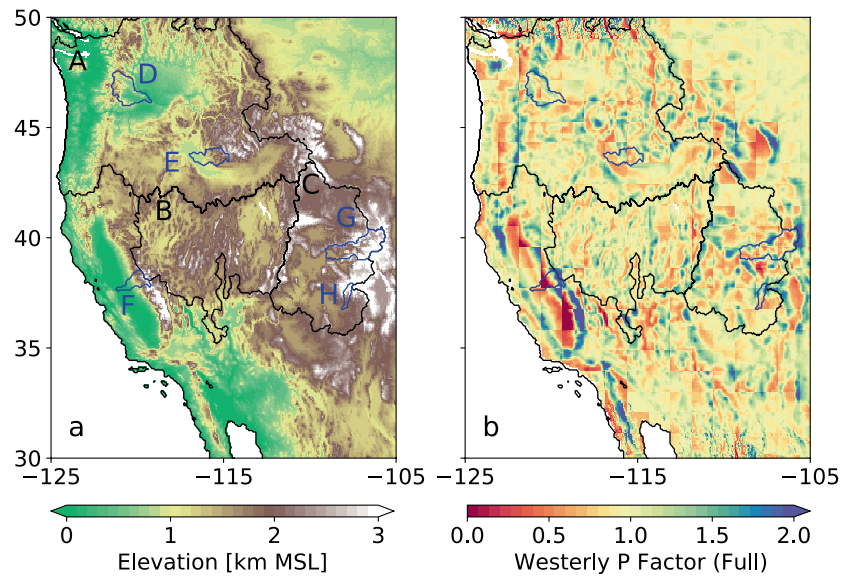
Figure 1. (a) Map of the domain and elevation (shading) used in the 1/16th-degree VIC model runs with select HUC2 regions outlined in black: (a) Pacific Northwest, (b) Great Basin, (c) Upper Colorado, and select HUC8/12 regions outlined in blue: (d) Yakima, (e) Boise, (f) Tuolumne, (g) Upper Colorado at Cameo, and (h) Animas. (b) The factor applied to P in the Full adjustment under Westerly winds. Visual artifacts in the North and South extremes of (b) are the result of the elevation data in these regions not being smoothed as the rest of the domain is and a magnifying effect of the Full adjustment.

3. Methods

Results focus on the contiguous United States (CONUS), West of the continental divide (bounded by 125°–105°W and 30°–50°N). Figure 1a shows the 1/16th-degree elevation of the domain. All results are also presented with respect to the hydrological water year, which runs from 1 October to 30 September, to ensure seasonally related maximum precipitation, streamflow, and snowmelt periods are grouped together. Black outlines in Figure 1 indicate 3 regions from the USGS Hydrologic Unit Code (HUC) level 2, and blue outlines indicate 5 smaller regions which are HUC levels 8 or 12 (Seaber et al., 1987). The modeled runoff from these regions is examined in Section 4. The HUC2 regions cover the major western basins, and are used to illustrate the effect of downscaling and bias correction at scales resolved by GCMs. The California HUC2 region is omitted because CESM2-Raw runoff already compared well to Livneh and the Full adjustment made only minor improvements (not shown). The Lower Colorado region is omitted since most of its streamflow originates in the Upper Colorado region. The smaller HUC8/12 basins are large enough to span multiple grid cells in CESM2, but small enough to illustrate the effect of local downscaling and bias correction without incorporating potentially compensating errors in the meteorological data.

This remainder of this section details the various adjustments applied to TMin, TMax, and P, from the most complex (Full-BC) to least complex (Elev) adjustment.

3.1. Full Adjustment With Bias Correction (Full-BC)

The seasonal bias corrections for TMin, TMax, and P closely follow Wood et al. (2002, 2004). Biases were computed from the seasonal averages of observations (aggregated to the CESM2 grid) and CESM2 output, using boreal seasons (December–February; March–May; June–August; and September–November). Precipitation bias was computed as a fraction of the CESM2 values, and temperature biases were computed as a difference, giving the following corrections:

$$\text{TMin}_{\text{BC}} = \text{CESM2(TMin)} - \left(\overline{\text{CESM2}}(\text{TMin}) - \overline{\text{Obs}}(\text{TMin}) \right), \tag{1}$$

$$\text{TMax}_{\text{BC}} = \text{CESM2(TMax)} - \left(\overline{\text{CESM2}}(\text{TMax}) - \overline{\text{Obs}}(\text{TMax}) \right), \tag{2}$$

$$P_{BC} = CESM2(P) \times \left(\frac{\overline{Obs(P)}}{\overline{CESM2(P)}} \right). \tag{3}$$

The fully adjusted and bias-corrected temperatures are derived by applying lapse rates (Γ) to the bias-corrected TMin and TMax to mimic the quasi-adiabatic warming and cooling associated with elevation changes:

$$TMin_{Full-BC} = TMin_{BC} - \Gamma_{TMin}\Delta z, \tag{4}$$

$$TMax_{Full-BC} = TMax_{BC} - \Gamma_{TMax}\Delta z. \tag{5}$$

Where Δz represents the 1/16th-degree elevation minus the CESM2 elevation at the gridpoint of interest (in km), and Γ_{TMax} and Γ_{TMin} are lapse rates. Derivation of these lapse rates from the Livneh data is described in Section 3.5. On occasion the use of different values for Γ_{TMin} and Γ_{TMax} results in TMax $\leq$ TMin. In this case, the two downscaled values are swapped.

In addition to Δz, the precipitation downscaling incorporates upwind topographical features to capture the enhanced and diminished precipitation on windward and leeward sides of mountains. To support these terms, look-up tables were generated containing the highest and lowest elevation within 100 km and 11.25° of 16 compass points (i.e., N, NNE, NE, ENE, etc.) for every 1/16th-degree gridpoint. The upwind direction is computed from the 700-mb CESM2 winds at the gridpoint of interest and rounded to the nearest of the 16 compass points. The minimum and maximum elevation differences, $\Delta z_{min,\ 100}$ and $\Delta z_{max,\ 100}$, are calculated from the minimum and maximum elevations within 100 km upwind minus the elevation of the gridpoint under consideration, all using the 1/16th-degree resolution elevation data. Similarly, $\Delta z_{min,\ 50}$ is the minimum elevation within 50 km upwind minus the elevation at the gridpoint. A further term, $\Delta z_{max,\ 50}$ had little impact on the results (not shown). The general form for the Full-BC precipitation is:

$$P_{Full-BC} = \max\left(0, P_{BC} \times \left[1 - a\Delta z_{max,100} + b\Delta z_{min,100} - c\Delta z_{min,50} + d\Delta z\right]\right). \tag{6}$$

3.2. Full Mass-Conserving Downscaling (Full)

The Full (mass-conserving) adjustments for temperature are absolute-value-conserving and use the non-bias-corrected TMin and TMax. The downscaled values are then scaled to keep average TMin and TMax over every CESM2 grid box unchanged. The Full adjustment for precipitation mimics Equation 6, except the non-bias-corrected precipitation from CESM2 is used, and the results are scaled to maintain consistent total precipitation over every CESM2 grid cell. The mass-conserving factor applied to P in the Full adjustment under westerly winds is shown in Figure 1b to demonstrate the adjustment moves precipitation from lower to higher elevations and from leeward to windward mountain sides.

3.3. Elevation-Only Mass-Conserving Downscaling (Elev)

The Elev adjustment uses the same temperature as the Full adjustment, but includes only the $d\Delta z$ term in the adjustment to the non-bias-corrected precipitation from CESM2:

$$P_{Elev} = \max(0, P \times [1 + d\Delta z]). \tag{7}$$

Like in the Full adjustment, the resulting precipitation is scaled to maintain consistent total precipitation over every CESM2 grid cell.

3.4. Tuning the Coefficients

Annual averages of bias-corrected variables from CESM2 and observations were compared using linear regressions to derive lapse rates and coefficients for Equations 4–7. After omitting points where observations and CESM2 differed by <0.1 K, comparison of the 1960–2009 mean annual TMin and TMax for each CESM2 grid cell returned coefficients of $\Gamma_{TMin} = 6.91$ K km^{-1} and $\Gamma_{TMax} = 6.65$ K km^{-1}. All 10 ensemble members used for training (1–10) were included in the regression so each 1/16th-degree observation was compared to Δz

and 10 CESM2 values. The statistically derived coefficients are close to the canonical environmental lapse rate of 6.5 K km^{-1}, lending confidence to their physical plausibility.

Precipitation coefficients $a = 0.199$, $b = 0.216$, $c = 0.675$, and $d = 0.557$ km^{-1} were obtained using a similar linear regression for P_{BC}, omitting points where observations and CESM2 differed by less than 10% of the CESM2 precipitation. Δz is constant for every observed grid point, but the other predictands are not since they depend on the 700-mb CESM2 wind direction. For this reason, the annual means of $\Delta z_{max,\ 100}$, $\Delta z_{min,\ 100}$, and $\Delta z_{min,\ 50}$, weighted by P_{BC}, were used in the regression. Similar to temperature, the use of 10 ensemble members meant each observational grid point was compared to 10 sets of P_{BC} and predictands.

3.5. VIC Simulations and Comparisons

VIC (Liang et al., 1994) version 5 (Hamman et al., 2018) was forced using TMin, TMax, and P, from (a) the Livneh data set (Livneh); (b) CESM2 output regridded to the observational grid using nearest-neighbor resampling (CESM2-Raw); (c) CESM2 output with the Elev adjustment applied (CESM2-Elev); (d) CESM2 output with the Full adjustment applied (CESM2-Full); and (e) CESM output with the Full-BC technique applied (CESM2-Full-BC). All five simulations used climatological winds (Livneh et al., 2013) averaged for the day of the year. TMin, TMax, P, and windspeeds were disaggregated to 3-hourly estimates using the MetSim software package (Bennett et al., 2020a), using the option to disaggregate precipitation evenly throughout the day. VIC was initialized on 1 January 1960 and 30-year averages are presented for water years 1980–2009 (1 October 1979–30 September 2009). Aside from TMin, TMax, and P forcings, all VIC simulations were configured identically; see Open Research for a sample parameter file.

4. Results

Figure 2a–2e show the mean annual precipitation from 1980 to 2009 for the various P forcings, while Figure 2f–2i show the deviation from Livneh. Similarly, mean annual runoff from the VIC runs are shown in Figure 2i–2n with deviations from Livneh in Figure 2o–2r. The CESM2-Full-BC precipitation and runoff are the most similar to Livneh, demonstrating the potential of the Full adjustment if bias correction can be achieved through other means. While not as close to Livneh as the CESM2-Full-BC output, the CESM2-Full output captures the lows and highs in precipitation and runoff seen in the Livneh panels better than the CESM2-Raw and CESM2-Elev output. This is especially noticeable along the Coastal and Cascade ranges where the distinction between the two ranges and the valleys between becomes more pronounced and similar to Livneh moving from left to right from CESM2-Raw to CESM2-Elev, CESM2-Full, and finally CESM2-Full-BC. A similar pattern of progressively closer matches to Livneh can be seen in the hydrographs for the eight subregions outlined in Figures 1 and 2 (Figure 3).

Mean annual hydrographs are shown in Figure 3 for the outlined regions. These hydrographs have been smoothed using a 7-day moving average to increase legibility of the plots. The runoff centroid timing (CT; e.g., Maurer et al., 2007), defined as the day of the water year where the cumulative annual runoff exceeds 50% of the total annual runoff (surface runoff plus baseflow), is marked with a vertical line. The CT and root mean squared error (RMSE) of the simulated mean daily runoff are used to quantify how well the different hydrographs match Livneh and are listed in Table S1. Both RMSE and CT are computed using the 7-day smoothed data to reduce sensitivity to short-term weather variability.

For most of the watersheds shown in Figure 3, CESM2-Full-BC (red) compares best to Livneh (i.e., smallest RMSE and difference in CT). Despite the bias-corrected TMin, TMax, and P, the CESM2-Full-BC still produces biased runoff in some basins (e.g., Yakima, Tuolomne, Boise, and Pacific Northwest). This is expected since TMin, TMax, and P were bias-corrected for each CESM2 grid cell, and not for the specific 1/16th degree grid-cells simulated. This is one limitation of the generalized method used.

While CESM2-Full-BC generally matches Livneh best, CESM2-Full is also effective at reducing RMSE for 6 out of 8 regions and in bringing CT closer to Livneh in all regions. The two regions where CESM2-Full increased RMSE compared to CESM2-Raw are the Great Basin and Upper Colorado, which are both HUC2 regions with wet biases in the CESM2-Raw precipitation. This suggests that for large areas with pervasive precipitation biases, correction of those biases is a more critical issue than resolving finer scale topographical features.
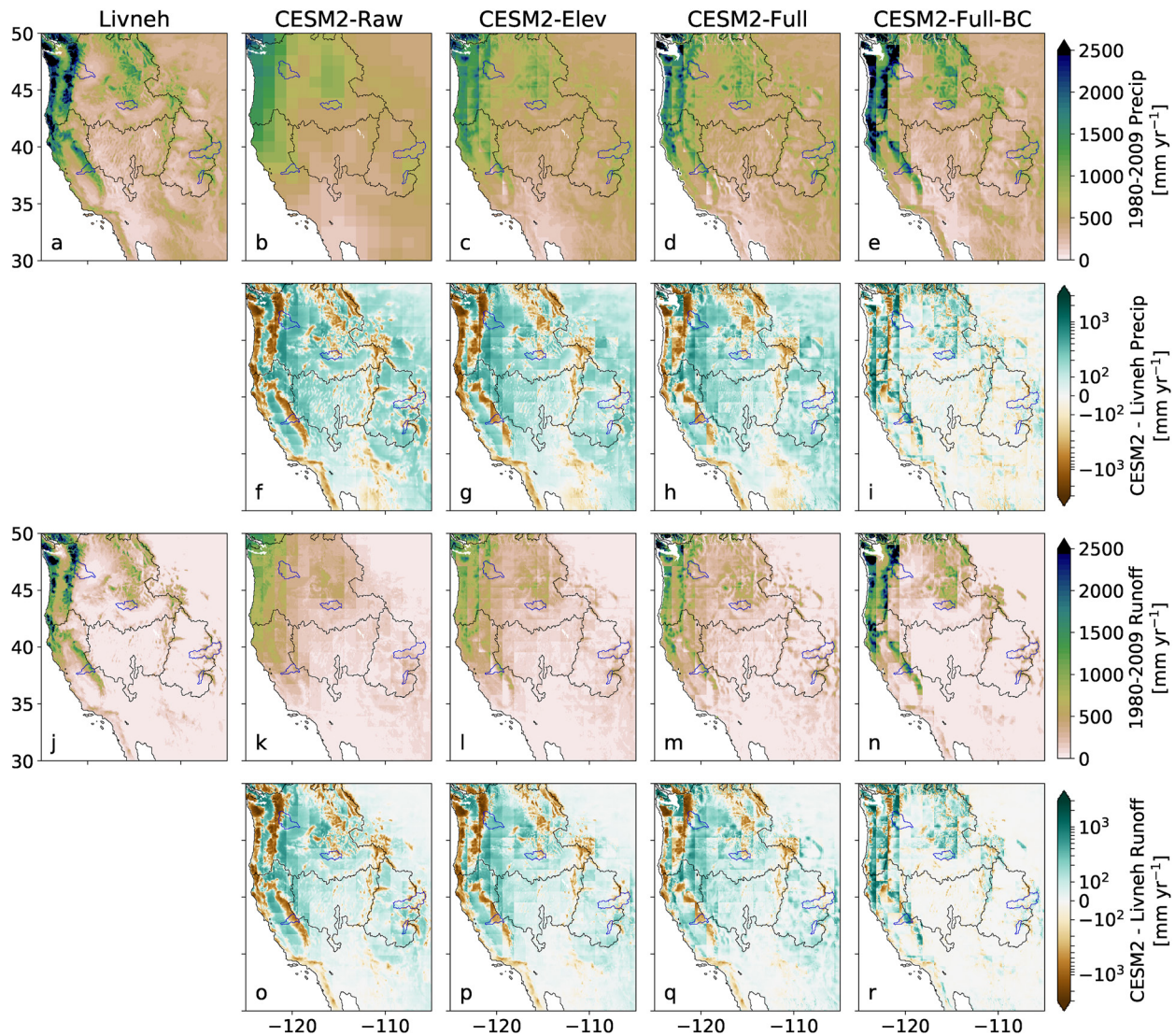
Figure 2. (a–e) Mean annual precipitation, (j–n) runoff from VIC, and (f–i, o–r) differences thereof from Livneh from water years 1980–2009 for: (a, j) Livneh, (b, f, k, and o) CESM2-Raw, (c, g, l, and p) CESM2-Elev, (d, h, m, and q) CESM2-Full, and (e, i, n, and r) CESM2-Full-BC.

CESM2-Full and CESM2-Elev give very similar results, though CESM2-Full results in CT's up to 4 days closer to Livneh. From the vertical lines in Figure 3 and Table S1, it is evident that the CT's from CESM2-Full match the Livneh CTs at least as well as those from CESM2-Elev. The RMSE for CESM2-Elev is slightly lower than CESM2-Full in some of the basins, however (e.g., Upper Colorado, Boise, and Animas). This demonstrates the wind-direction-dependent terms in the Full correction improve the prediction of runoff centroid timing, but add little value to the prediction of the overall hydrograph shape beyond that of the elevation-only adjustment.

5. Discussion

5.1. Value for Prediction of Runoff Timing

The main benefit of the Full technique is that the adjusted precipitation and temperature produce runoff CTs closer to those estimated from historical precipitation and temperature, as simulated by VIC, than the raw CESM2 output. The Full technique also slightly improves upon the Elev adjustment. This suggests that Full adjustment could increase the reliability of future projections of CT if coupled with sub-grid-scale hydrological components of GCMs. This would be beneficial for water resource management since timing is important for planning
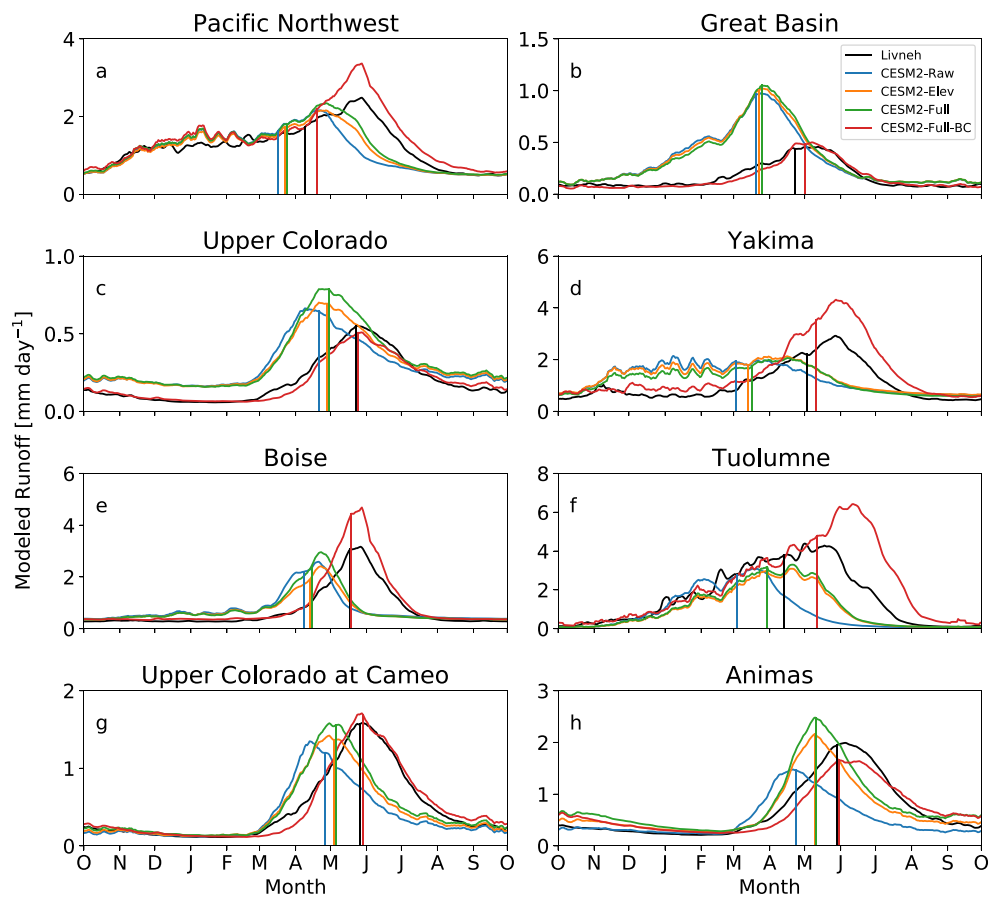
Figure 3. Mean annual cycle of runoff averaged over hydrological regions for water years 1980–2009 for: Livneh (black), CESM2-Raw (blue), CESM2-Elev (orange), CESM2-Full (green), and CESM2-Full-BC (red). Vertical colored lines mark the CT. Table S1 lists the RMSEs and error in CT for each colored hydrograph compared to Livneh.

the storage and release of reservoirs, especially where reservoir capacity may not be sufficient to store earlier melt-water for later use, as is the case in much of the western US (Barnett et al., 2005).

Compared to other downscaling and bias-correction methods, ours has the advantage of being computationally inexpensive and easily applied to large ensembles of GCM simulations. This is in contrast to dynamical downscaling and regional climate models which may compare better to observations but require much more extensive computing power (e.g., Chan et al., 2012; Mearns et al., 2009; Prein et al., 2013, 2015; Solman et al., 2021). Our method is also less dependent on observations than many statistical downscaling techniques since it is not tuned to individual grid cells (e.g., Gutmann et al., 2022; Maurer & Hidalgo, 2008; Pierce et al., 2014; Wood et al., 2004). For example, Pierce et al. (2023) produced CMIP6 model output downscaled to 6 km over much of North America using localized constructed analogs. The result was a valuable data set for examining North American climate projections, but the downscaling method cannot be applied to data-sparse regions due to the extensive observational data set needed to produce the analogs.

While the coefficients used in all three adjustments were derived using observations and model output from the western US, the method is based in the physics of orographically driven precipitation and quasi-adiabatic temperature changes with altitude. This physical basis, plus the derivation of coefficients using the entire contiguous United States West of the continental divide—rather than a specific watershed or individual grid points—makes it possible to apply the adjustment globally. This generalizability is also supported by the variety of climates represented in the western contiguous United States. According to the Köppen-Geiger system (Peel et al., 2007), these climates range from hot desert to cold climates with no dry season, mainly lacking in tropical and polar climates.

5.2. Uncertainties and Future Evaluation

There are several sources of uncertainty in the analysis conducted, including the Livneh data set, the VIC model, and the use of a single ensemble member for verification. The Livneh data set is not purely observational, but is a processed data set where point observations were interpolated to a gridded product. This introduces errors and uncertainties regarding the true precipitation and temperature. The disaggregation of daily observations and model output to 3-hourly input for VIC introduces additional uncertainty, especially considering the rate of precipitation can have a large impact on the proportional amount of runoff given the same total precipitation. Use of the VIC model to simulate runoff, along with uncertainties in the parameters (e.g., land use/coverage, soil types) and the lack of river routing also limits the utility of the results to water resource managers. The lack of river routing could also explain why our results differ from those from a recent study in the Mekong River Basin which found that higher resolution precipitation data did not improve hydrological model performance (Kabir et al., 2022). Differences in regional hydrology and geography—especially the proportion of runoff originating from snowment—and/or the hydrological model used may also explain this discrepancy, however. Expanded studies including river routing and using multiple ensemble members would help bound the uncertainty in the utility of our method. The utility of our adjustment is also predicated on the ability of GCMs to predict long-term changes and variability on a larger scale demonstrated by previous studies (e.g., Danabasoglu et al., 2020; Hausfather et al., 2020; Phillips et al., 2020; Simpson et al., 2020).

In addition to expanding our analysis to include river routing and additional ensemble members, water resource managers would benefit from an analysis of the impacts of climate change on the frequency of extreme swings between dry and wet years relative to historical experiences (e.g., Morss et al., 2018; Prendergrass et al., 2017; Swain et al., 2018; Tye et al., 2023). An analysis of high and low flows at the daily timescale would also be helpful in establishing the Full adjustment's value to water resource planning and management in the context of extreme flow hazards.

5.3. Limitations and Possible Refinements

Despite the advantages over more locally tuned statistical methods, caution is warranted in applying our adjustments to other geographic areas since the appropriateness of the coefficients (a, b, c, and d in Equation 6) would be unknown. It would be advantageous to derive the coefficients for Equations 4–7 using global observations, possibly with different coefficients for different regions. While such observations are not readily available, very-high-resolution model simulations may be sufficient (e.g., Lundquist et al., 2019) to further improve generalizability. Such simulations could also aid in better capturing more localized relationships between topography, wind patterns, and precipitation which are not resolved with the roughly 100-km grid spacing of the 700 mb winds used herein. In its current state, the method could at least be used as one member in an ensemble of downscaling or bias correction methods to bound the uncertainty of climate projections in data-sparse regions.

The Full adjustment brings the hydrographs closer to those resulting from simulations using observations in some smaller (on the order of 100,000 km^2), snow-dominated basins (e.g., Upper Colorado at Cameo, Animas, Toulomne), as judged from RSME in the mean annual hydrograph. Results from CESM2-Full-BC also suggest that further reduction in RMSE may require correcting large-scale precipitation biases. This is especially true for the Great Basin and Upper Colorado HUC2s where the CESM2 has a large wet bias, particularly in the first half of the water year (Figures 2a, 2b, 3b, and 3c). One possible method to make further improvements to the hydrograph shape while maintaining mass conservation compatibility with GCM coupling would be through conservation of the total-atmosphere water column. This would mean converting atmospheric water vapor to/ from precipitation to alter the mass of precipitation, but maintaining the total mass of water. The incorporation of 700-mb wind speed, and more localized fitting to the Full adjustments are also being considered for future testing.

6. Conclusion

We presented a mass-conserving physically based method to downscale precipitation and temperature from CESM2 over mountainous terrain. The Full method uses the CESM2 modeled 700-mb wind direction and sub-grid-scale topographical features to obtain more realistic precipitation and temperature variables. This adjustment was compared to the Elev version which omitted upwind topographical features, and the non-mass-conserving Full-BC version which included a CESM2-grid-scale bias correction.

When VIC was forced using the adjusted variables, the Full-BC adjustment generally compared best to VIC output forced with observations. This bias correction was especially important for the large HUC2 basins of the Upper Colorado and Great Basin where the CESM2 precipitation bias is large and pervasive. The mass-conserving Full method still compared better to VIC output using observations than simulations using the raw CESM2 output or the Elev adjustment. The improvement was most noticeable in the sub-grid-scale variability in runoff over the Coastal and Cascade mountains, and the runoff CT of all basins examined. Additional evaluation of the adjustment on extreme high and low flows and other geographic areas would be beneficial future work.

The Full adjustment presented herein has the potential to improve the utility of GCM output to water resource management. This can be achieved by applying the method outside the GCM using a standalone hydrological model, as in this study. Since the method is mass-conserving, however, future work may allow for a fully coupled version to be integrated into the land components of GCMs. Such an integration would allow other components of the GCM (e.g., the atmospheric component) to benefit from the more accurate representation of surface hydrology.

Data Availability Statement

All data generated for this study (e.g., CESM2 and VIC output) along with a Jupyter notebook to recreate all tables and figures are available in a repository with the https://doi.org/10.5065/2tdq-fe83 (Rugg et al., 2023). The repository also contains parameter files for running MetSim and VIC.

Version 2.2.1 of MetSim used disaggregate CESM2 output prior to running VIC is preserved with https://doi.org/10.5281/zenodo.3728015 with open access (Bennett et al., 2020b).

Version 5.0.1 of VIC used to compute runoff from precipitation is preserved at https://doi.org/10.5281/zenodo.267178 with open access (Hamman et al., 2017).

References

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. https://doi.org/10.1038/nature04141

Bennett, A. R., Hamman, J. J., & Nijssen, B. (2020a). MetSim: A Python package for estimation and disaggregation of meteorological data. *Journal of Open Source Software*, *5*(47), 2042. https://doi.org/10.21105/joss.02042

Bennett, A. R., Hamman, J. J., & Nijssen, B. (2020b). MetSim: A Python package for estimation and disaggregation of meteorological data (v2.2.1) [Software]. Zenodo. https://doi.org/10.5281/zenodo.3728015

Chan, S. C., Kendon, E. J., Fowler, H. J., Blenkinsop, S., Ferro, C. A. T., & Stephenson, D. B. (2012). Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate Dynamics*, *41*(5–6), 1475–1495. https://doi.org/10.1007/s00382-012-1568-9

Collier, C. G. (1975). A representation of the effects of topography on surface rainfall within moving baroclinic disturbances. *Quarterly Journal of the Royal Meteorological Society*, *101*(429), 407–422. https://doi.org/10.1002/qj.49710142902

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Ekström, M., Gutmann, E. D., Wilby, R. L., Tye, M. R., & Kirono, D. G. C. (2018). Robustness of hydroclimate metrics for climate change impact research. *WIREs Water*, *5*(4), e1288. https://doi.org/10.1002/wat2.1288

Ghan, S. J., & Shippert, T. (2006). Physically based global downscaling: Climate change projections for a full century. *Journal of Climate*, *19*(9), 1589–1604. https://doi.org/10.1175/jcli3701.1

Gutmann, E. D., Hamman, J. J., Clark, M. P., Eidhammer, T., Wood, A. W., & Arnold, J. R. (2022). En-GARD: A statistical downscaling framework to produce and test large ensembles of climate projections. *Journal of Hydrometeorology*, *23*(10), 1545–1561. https://doi.org/10.1175/JHM-D-21-0142.1

Hamman, J., Nijssen, B., Bohn, T., Franssen, W., Yixinmao, M., Gergel, D., et al. (2017). UW-Hydro/VIC: VIC 5.0.1 (VIC.5.0.1) [Software]. Zenodo. https://doi.org/10.5281/zenodo.267178

Hamman, J. J., Nijssen, B., Bohn, T. J., Gergel, D. R., & Mao, Y. (2018). The variable infiltration capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. *Geoscientific Model Development*, *11*(8), 3481–3496. https://doi.org/10.5194/gmd-11-3481-2018

Hausfather, Z., Drake, H. F., Abbott, T., & Schmidt, G. A. (2020). Evaluating the performance of past climate model projections. *Geophysical Research Letters*, *47*(1), e2019GL085378. https://doi.org/10.1029/2019GL085378

Hobbs, P. V., Easter, R. C., & Fraser, A. B. (1973). A theoretical study of the flow of air and fallout of solid precipitation over mountainous terrain: Part II. Microphysics. *Journal of Atmospheric Sciences*, *30*(5), 813–823. https://doi.org/10.1175/1520-0469(1973)030<0813:ATSOTF>2.0.CO;2

Kabir, T., Pokhrel, Y., & Felfelani, F. (2022). On the precipitation-induced uncertainties in process-based hydrological modeling in the Mekong River Basin. *Water Resources Research*, *58*(2), e2021WR030828. https://doi.org/10.1029/2021WR030828

Lehner, F., Wood, A. D., Vano, J. A., Lawrence, D. M., Clark, M. P., & Mankin, J. S. (2019). The potential to reduce uncertainty in regional runoff projections from climate models. *Nature Climate Change*, *9*(12), 926–933. https://doi.org/10.1038/s41558-019-0639-x

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, *44*(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, *99*(D7), 14415–14428. https://doi.org/10.1029/94JD00483

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., et al. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *Journal of Climate*, *26*(23), 9384–9392. https://doi.org/10.1175/JCLI-D-12-00508.1

Lundquist, J., Hughes, M., Gutmann, E., & Kapnick, S. (2019). Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. *Bulletin of the American Meteorological Society*, *100*(12), 2471–2490. https://doi.org/10.1175/BAMS-D-19-0001.1

Maurer, E. P., & Hidalgo, H. G. (2008). Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, *12*(2), 551–563. https://doi.org/10.5194/hess-12-551-2008

Maurer, E. P., Stewart, I. T., Bonfils, C., Duffy, P. B., & Cayan, D. (2007). Detection, attribution, and sensitivity of trends toward earlier streamflow in the Sierra Nevada. *Journal of Geophysical Research*, *112*(D11), D11118. https://doi.org/10.1029/2006JD008088

Mearns, L. O., Gutowski, W. J., Jones, R., Leung, L.-Y., McGinnis, S., Nunes, A. M. B., & Qian, Y. (2009). A regional climate change assessment program for North America. *Eos*, *90*(36), 311–312. https://doi.org/10.1029/2009EO360002

Morss, R. E., Done, J. M., Lazrus, H., Towler, E., & Tye, M. R. (2018). Assessing and communicating uncertainty in decadal climate predictions: Connecting predictive capacity to stakeholder needs. *US CLIVAR Variations*, *16*(3), 24–30. https://doi.org/10.5065/D62N513R

Null, S. E., Viers, J. H., & Mount, J. F. (2010). Hydrologic response and watershed sensitivity to climate warming in California's sierra Nevada. *PLoS One*, *5*(4), e9932. https://doi.org/10.1371/journal.pone.0009932

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Phillips, A. S., Deser, C., Fasullo, J., Schneider, D. P., & Simpson, I. R. (2020). Assessing climate variability and change in model large ensembles: A user's guide to the "climate variability diagnostics package for large ensembles" version 1.0. https://doi.org/10.5065/h7c7-f961

Pierce, D. W., Cayan, D. R., Feldman, D. R., & Risser, M. D. (2023). Future increases in North American extreme precipitation in CMIP6 downscaled with LOCA. *Journal of Hydrometeorology*, *24*(5), 951–975. https://doi.org/10.1175/JHM-D-22-0194.1

Pierce, D. W., Cayan, D. R., & Thrasher, B. L. (2014). Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, *15*(6), 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., & Georgievski, G. (2013). Added value of convective permitting seasonal simulations. *Climate Dynamics*, *41*(9–10), 2655–2677. https://doi.org/10.1007/s00382-013-1744-6

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Georgen, K., et al. (2015). A review on regional convective-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*(2), 323–361. https://doi.org/10.1002/2014RG000475

Prendergrass, A. G., Knutti, R., Lehner, F., Deser, C., & Sanderson, B. M. (2017). Precipitation variability increases in a warmer climate. *Scientific Reports*, *7*(1), 17966. https://doi.org/10.1038/s41598-017-17966-y

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., et al. (2021). Ubiquity of human-induced changes in climate variability. *Earth Systems Dynamics*, *12*(4), 1393–1411. https://doi.org/10.5194/esd-12-1393-2021

Rugg, A., McCrary, R. R., & Gutmann, E. D. (2023). Mass-conserving downscaling of climate model precipitation over mountainous terrain for water resource applications. Version 1.0 [Dataset]. UCAR/NCAR - GDEX. Retrieved from https://gdex.ucar.edu/dataset/383_stebbins.html

Salathé, E. P., Jr. (2003). Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *International Journal of Climatology*, *23*(8), 887–901. https://doi.org/10.1002/joc.922

Schipper, J. W., Frueh, B., Pfeiffer, A., & Zaengl, G. (2011). Wind direction-dependent statistical downscaling of precipitation applied to the Upper Danube catchment. *International Journal of Climatology*, *31*(4), 578–591. https://doi.org/10.1002/joc.2084

Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). Hydrologic unit maps, U.S. *Geological Survey, Water Supply Paper*, *2294*, 63.

Sedláček, J., & Knutti, R. (2014). Half the world's population experience robust changes in the water cycle for a 2°C warmer world. *Environmental Research Letters*, *9*(4), 044008. https://doi.org/10.1088/1748-9326/9/4/044008

Simpson, I. R., Bacmeister, J., Neale, R. B., Hannay, C., Gettelman, A., Garcia, R. R., et al. (2020). An evaluation of the large-scale atmospheric circulation and its variability in CESM2 and other CMIP models. *Journal of Geophysical Research: Atmospheres*, *125*(13), e2020JD032835. https://doi.org/10.1029/2020JD032835

Smith, R. B., & Barstad, I. (2004). A linear theory of orographic precipitation. *Journal of the Atmospheric Sciences*, *61*(12), 1377–1391. https://doi.org/10.1175/1520-0469(2004)061<1377:ALTOOP>2.0.CO;2

Solman, S., Jacob, D., Frigon, A., Teichmann, C., Rixen, M., Gutowski, W., & Lake, I. (2021). The future scientific challenges for CORDEX. Retrieved from https://cordex.org/wp-content/uploads/2022/08/The-future-of-CORDEX-MAY-17-2021_2.pdf

Swain, D. L., Langenbrunner, B., Neelin, J. D., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, *8*(5), 427–433. https://doi.org/10.1038/s41558-018-0140-y

Takayabu, I., Kanamaru, H., Dairaku, K., Benestad, R., von Storch, H., & Christensen, J. H. (2016). Reconsidering the quality and utility of downscaling. *Journal of the Meteorological Society of Japan*, *94A*(0), 31–45. https://doi.org/10.2151/jmsj.2015-042

Tesfa, T. K., Leung, L. R., & Ghan, S. J. (2020). Exploring topography-based methods for downscaling subgrid precipitation for use in Earth System Models. *Journal of Geophysical Research: Atmospheres*, *125*(5), e2019JD031456. https://doi.org/10.1029/2019JD031456

Tye, M., Richter, J., Wieder, W., Wood, A., Gutmann, E., Haupt, S., et al. (2023). Water availability metrics workshop report. *Open Science Framework*. https://doi.org/10.17605/OSF.IO/M7NXD

Viviroli, D., Archer, D. R., Buytaert, W., Fowler, H. J., Greenwood, G. B., Hamlet, A. F., et al. (2011). Climate change and mountain water resources: Overview and recommendations for research, management, and policy. *Hydrology and Earth System Sciences*, *15*(2), 471–504. https://doi.org/10.5194/hess-15-471-2011

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2006). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*(7), W07447. https://doi.org/10.1029/2006WR005653

Viviroli, D., Kummu, M., Meybeck, M., Kallio, M., & Wada, Y. (2020). Increase dependence of lowland populations on mountain water resources. *Nature Sustainability*, *3*(11), 917–928. https://doi.org/10.1038/s41893-020-0559-9

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrological implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, *62*(1–3), 182–216. https://doi.org/10.1023/B:CLIM.0000013685.99609.9e

Wood, A. W., Maurer, E. P., Kuman, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research*, *107*(D20), 4429. https://doi.org/10.1029/2001JD000659