Geologic controls on apparent root-zone storage capacity

W. Jesse Hahm[1], David N Dralle[2], Dana A Lapides[3], Robert Ehlert[1], and Daniella Rempe[4]

[1]Simon Fraser University
[2]Pacific Southwest Research Station, United States Forest Service
[3]Pacific Southwest Research Station, United States Forest Service — Simon Fraser University
[4]University of Texas at Austinn

May 25, 2023

Abstract

The water storage capacity of the root zone determines whether plants survive dry periods and controls the partitioning of precipitation into streamflow and evapotranspiration. It is currently thought that top-down, climatic factors are the primary control on this capacity via their interaction with plant rooting adaptations. However, it remains unclear to what extent bottom-up, geologic factors can provide an additional constraint on storage capacity. Here we use a machine learning approach to identify regions with lower than climatically expected apparent storage capacity. We find that in seasonally dry California these regions overlap with particular geologic substrates. We hypothesize that these patterns reflect diverse mechanisms by which substrate can limit storage capacity, and highlight case studies consistent with limited weathered bedrock extent (melange in the Northern Coast Range), toxicity (ultramafic substrates in the Klamath-Siskiyou region), nutrient limitation (phosphorus-poor plutons in the southern Sierra Nevada), and low porosity capable of retaining water (volcanic formations in the southern Cascades). The observation that at regional scales climate alone does not 'size' the root zone has implications for the parameterization of storage capacity in models of plant dynamics (and the interrelated carbon and water cycles), and also underscores the importance of geology in considerations of climate-change induced biome migration and habitat suitability.

Geologic controls on apparent root-zone storage capacity

W.J. Hahm[1], D.N. Dralle[2], D.A. Lapides[1,2], R.S. Ehlert[1], D.M. Rempe[3]

[1]Simon Fraser University, Burnaby, BC, Canada

[2]Pacific Southwest Research Station, United States Forest Service, Davis, CA, USA

[3]University of Texas at Austin, Austin, TX, USA

Key Points:

- Regionally extensive areas of low apparent root-zone storage capacity for a particular climate coincide with particular geologic substrates
- Hypothesized geologic controls include water storage capacity limitation, nutrient limitation, and toxicity

Corresponding author: WJ Hahm, `whahm@sfu.ca`

Abstract

The water storage capacity of the root zone determines whether plants survive dry periods and controls the partitioning of precipitation into streamflow and evapotranspiration. It is currently thought that top-down, climatic factors are the primary control on this capacity via their interaction with plant rooting adaptations. However, it remains unclear to what extent bottom-up, geologic factors can provide an additional constraint on storage capacity. Here we use a machine learning approach to identify regions with lower than climatically expected apparent storage capacity. We find that in seasonally dry California these regions overlap with particular geologic substrates. We hypothesize that these patterns reflect diverse mechanisms by which substrate can limit storage capacity, and highlight case studies consistent with limited weathered bedrock extent (melange in the Northern Coast Range), toxicity (ultramafic substrates in the Klamath-Siskiyou region), nutrient limitation (phosphorus-poor plutons in the southern Sierra Nevada), and low porosity capable of retaining water (volcanic formations in the southern Cascades). The observation that at regional scales climate alone does not 'size' the root zone has implications for the parameterization of storage capacity in models of plant dynamics (and the interrelated carbon and water cycles), and also underscores the importance of geology in considerations of climate-change induced biome migration and habitat suitability.

Plain Language Summary

What determines how much water plants can store in their root zone? One school of thought posits that plants 'size' the root-zone capacity to survive a drought of a particular return period. In this scenario, plants extend their roots into the subsurface in response to climate drivers (e.g., precipitation magnitude-frequency and atmospheric water demand). This worldview neglects the potential for geology to restrict root access to water. 'Bottom-up' limitations on storage capacity have been described at individual field sites, but it has been unclear how to identify geologic limitations at large scales. Here, we introduce an approach that quantifies differences between the climatically expected and locally observed apparent storage capacity, and relate these spatial patterns to geologic substrate. Importantly, we quantify apparent storage capacity via a method that includes water below the upper 1.5 m, within weathered bedrock, which is an important water source in seasonally dry climates and is typically excluded from traditional soil texture databases. We find that geology limits storage capacity at regional scales, and synthesize existing field evidence to hypothesize mechanisms of bottom-up control. Our findings have important implications for water-carbon cycle modeling efforts and the prediction of plant biome migration in response to climate change.

1 Introduction

Root-accessible water storage capacity in the subsurface is a key earth system property that regulates the water and carbon cycles (Kleidon & Heimann, 1998). For example, plant transpiration of stored water is a first-order control on Earth's surface energy budget and terrestrial water partitioning (Milly, 1994), setting aquatic ecosystem habitat and water quality and quantity for downstream users. Sufficient storage capacity also enables plants to bridge meteorologic droughts and sustain photosynthesis during extreme dry periods (Porporato et al., 2004; McLaughlin et al., 2020). It has been argued that top-down (climatic) drivers are primarily responsible for determining the large-scale spatiotemporal variability of storage capacity (Nijzink et al., 2016; Guswa, 2008, 2010; M. Liu et al., 2022; van Oorschot et al., 2021; Bouaziz et al., 2022). However, field investigations have revealed that geologic or edaphic factors can exert a primary control at some sites (e.g., Hahm et al., 2019), but it is presently unknown where and why geologic factors eclipse climate factors at landscape scales. This uncertainty challenges earth system and dynamic global vegetation modeling efforts, including prediction of plant biome migration in the context of climate change.

Plant-available water storage capacity is understood to be set by i) the porosity profile, which determines the amount of water that can be held at various water potentials, and ii) the presence of roots to access that porosity (Klos et al., 2018; C. Zhang et al., 2020). Factors related to geology can limit the storage capacity, either directly (by restricting accessible porosity in the near surface (e.g, the presence of low porosity fresh bedrock at a shallow depth) (Hahm et al., 2019) or by being too permeable to store water (H. Liu et al., 2021; Jiang et al., 2020)) or indirectly (by inhibiting plant growth (via nutrient limitation or toxicity) (Hahm et al., 2014; Kruckeberg, 1985; Morford et al., 2011) that in turn inhibits root exploitation of accessible porosity), as depicted in Figure 1. In contrast, top-down (climatic) controls are thought to determine the storage capacity primarily by setting atmospheric water demand and precipitation inputs, including the frequency and duration of dry periods that plants need to endure to survive. Various models that explore optimal plant strategy suggest that plants will invest just enough carbon in root profiles to have sufficient water access to survive dry periods of a particular recurrence interval (Schymanski et al., 2008; Schenk, 2008; Yang et al., 2016; Speich et al., 2018; Guswa, 2008, 2010). This school of thought is encapsulated in the notion that climate 'sizes' the root-zone storage capacity (Gao et al., 2014; de Boer-Euser et al., 2016, 2019; Gentine et al., 2012). Optimal rooting frameworks may neglect the potential for bottom-up factors to limit storage capacity, however, because they implicitly treat the subsurface like an infinite sand box, into which plants may invest as much—or as little—into rooting as is advantageous (e.g., Singh et al., 2020).

A first-order challenge in teasing apart climatic versus geologic controls on storage capacity is quantifying the actual storage capacity accessed by plants. Traditionally, the storage capacity has been parameterized in models through calibration or with the aid of distributed soils datasets, which typically quantify water retention properties through the upper 1 to 1.5 m or to the depth of a restrictive layer. Although widely available and relatively finely resolved, soils datasets have two principle shortcomings: i) they do not capture whether roots are actually present in the soil profile, and ii) they do not extend deep enough into the subsurface to capture porosity profiles in deeper weathered bedrock that commonly underlies soils (Holbrook et al., 2014; Witty et al., 2003; Dawson et al., 2020), where widespread evidence has emerged of root penetration and water uptake (McCormick et al., 2021; Zhu et al., 2023; Stocker et al., 2023). The relative inaccessibility of the deep subsurface challenges quantification of these factors (Stocker et al., 2023).

A recently developed and now widely adopted alternative approach (Wang-Erlandsson et al., 2016; Dralle et al., 2021) constrains storage capacity via tracking of hydrologic fluxes. Precipitation (flux in) and evapotranspiration (flux out) are monitored at a location, and it is reasoned that the root-accessible subsurface water storage capacity must be big enough to explain the largest observed cumulative evapotranspiration in excess of precipitation over a period of record (i.e., the largest observed storage deficit). This approach quantifies an *apparent* root-zone water storage capacity (S_R): i.e., S_R identified from the largest observed deficit is only a lower bound on actual accessible storage capacity (McCormick et al., 2021). For example, it is possible that plants may have had access to—and would have used—more water if dry conditions persisted. In other words, actual root-zone water storage capacity may be larger than S_R, but we do not have the means to directly measure it (although some researchers have attempted to quantify it by fitting yearly maximum deficit values to extreme value distributions (Wang-Erlandsson et al., 2016)). Nevertheless, storage capacity calculated via deficit-style approaches has many theoretical and pragmatic advantages. S_R results in improved hydrological model performance when used as an input parameter (Wang-Erlandsson et al., 2016) and can explain continental-scale patterns in water partitioning (Cheng et al., 2022) and storage dynamics (Trautmann et al., 2022); deficit calculations have also proven essential in the accurate prediction of snowmelt contributions to streamflow following droughts (Lapides et al., 2022). Importantly, deficit-calculated S_R does not require *a priori* assumptions regarding porosity or rooting profiles; distributed hydrologic flux datasets make it feasible to estimate S_R at large spatial scales in cloud-based analysis platforms like Google Earth Engine.

Although distributed estimates of S_R are now available, it has remained challenging to isolate both the spatial patterns and drivers of geologic factors impacting the magnitude of S_R. Here we explore a simple machine learning approach to predict S_R, assuming that climatic controls are the primary drivers of spatial variations in S_R. This exercise reveals locations where the null hypothesis may be rejected (i.e., places where geologic controls may be important) based on deviations between the S_R predicted by modern climate (informed by all observations) and empirically observed S_R (the local observation). We then explore select case studies of geologic control and suggest process explanations through an analysis of available subsurface geologic and hydrologic field studies.

2 Methods

2.1 Study area

The study area covers the state of California, USA, where three factors make for an ideal setting to explore geologic controls on S_R: i) there is a high diversity of annual precipitation and potential evapotranspiration rates, geologic substrates, and tectonic uplift rates, resulting in large spatial gradients to explore controls on plant biomes and S_R; ii) the local Mediterranean climate (asynchronous seasonal precipitation and energy input, with a long summer dry period) results in almost complete reliance on wet season-replenished storage to sustain evapotranspiration in summer, and iii) existing evidence for widespread and routine use of bedrock water by woody vegetation (McCormick et al., 2021) indicates that water storage capacity inferred from soils databases is insufficient to describe S_R and that bedrock geologic properties that impact plants (nutrients, toxins, and water status) are likely to strongly influence spatial patterns in S_R.

2.2 Identification of lower than climatically expected S_R

To identify locations with a geologic control on S_R, we compare observed S_R to climatically predicted S_R on a per-pixel basis. Locations with an observed S_R lower than expected for the local climate (i.e., low relative to the predicted S_R) are potentially indicative of a geologic limiting factor. The observed S_R is determined based on the previously described approach that records at each location the maximum deficit between cumulative precipitation and cumulative evapotranspiration (Wang-Erlandsson et al., 2016; Dralle et al., 2021), which in California typically exceeds published soils database water storage capacities and must include deeper water storage in bedrock (McCormick et al., 2021). We use a machine learning (random forest) model to predict S_R solely as a function of climatic factors.

2.3 Data sources

All datasets described below previously existed and were ingested and analyzed for this study via the Google Earth Engine cloud computation environment (Gorelick et al., 2017), where spatial joins and spatial resampling were also performed. The data are mapped at the state-wide level in Figure 2.

2.3.1 Observed apparent root-zone water storage capacity, S_R

S_R was calculated following the deficit-based approach described above (see Wang-Erlandsson et al. (2016) for more details), modified to account for the impacts of snow following Dralle et al. (2021). We used the S_R dataset provided by Dralle et al. (2021), which was calculated using data from 2003-2017 and is provided at approximately 1 km pixel resolution. This S_R dataset relies on precipitation data from PRISM (Daly et al., 2015), evapotranspiration data from PML v2 (Y. Zhang et al., 2019), and snow cover from the MODIS Terra normalized difference snow index product (Hall et al., 2010). This S_R dataset also excludes urban areas, open water, and croplands as well as areas in which evapotranspiration exceeded precipitation, which may be due to unaccounted for irrigation, inter-pixel groundwater fluxes or data error.

Consumer dynamics (Kuijper et al., 2015) or episodic disturbances (e.g., fire or logging) may result in lower than climatically possible evapotranspiration and therefore a lower than climatically expected S_R. This is particularly of concern when S_R is inferred from a relatively short timeseries of precipitation and evapotranspiration. Here, because the S_R dataset is inferred from 15 continuous water years, we do not exclude areas with logging or fire. This is motivated by the desire to include as much training data as possible and the finding that paired catchment studies in the region have observed non-detectable changes in dry season streamflow (the time of year when deficits accumulate) just five years after logging (Keppeler & Ziemer, 1990). Spot checking of logged areas indicates that S_R differences between adjacent logged or burned areas during the study period tend to be small relative to differences across geologic contacts or large climate zones.

2.3.2 Climatic predictors of S_R

We used four static climate variables as predictors of S_R:

- Mean annual precipitation, P (mm)
- Mean annual potential evapotranspiration PET (mm)
- The coefficient of variation of annual precipitation, CV_P, equal to the standard deviation of annual precipitation divided by mean annual precipitation.
- The asynchronicity index between precipitation and potential evapotranspiration (in time and in relative magnitude), ASI (Feng et al., 2019)

The precipitation data were obtained from PRISM (Daly et al., 2015) and the potential evapotranspiration data from the MODIS Terra MOD16A2 product (Running et al., 2017) for the period 2003-2017. The ASI raster was previously generated and described in (Hahm, Lapides, et al., 2022). These climate variables were chosen for their widespread availability at relatively high spatial resolution, and because magnitudes and timing of water delivery and water demand are the first order constraints on the amount of water available to plant biomes and the amount that can be taken up by the atmosphere; together P and PET also capture the aridity index (which is important for water partitioning within the classical Budyko framework). The variability of annual precipitation (captured in CV_P) roughly accounts for drought recurrence intervals, which have been hypothesized to be the other primary climatic driver of top-down root zone storage capacity.

2.3.3 Random forest model

We used the RandomForestRegressor module within the scikit-learn Python package (Pedregosa et al., 2011) to predict S_R from four climate variables (mean annual precipitation, the coefficient of variation of annual precipitation, mean annual potential evapotranspiration, and the seasonal asynchronicity between precipitation and energy delivery; detailed in Section 2.3.2). Model accuracy was assessed by first training on a random subset of 75% of the observations and using the resulting preliminary model to predict S_R with the remaining 25% set-aside validation data, after which a final model was trained on the entire dataset. In each case default scikit-learn (version 1.2.0) hyperparameters were used, except for the minimum number of samples per leaf node, which was set to 100 (discussed below).

The choice of random forest modeling over a multiple linear regression approach (with and without interaction terms) is due to the flexibility of the random forest to account for non-linear interactions between climate drivers and S_R, which were were apparent during exploratory data analysis. The choice of random forest modeling over climatic envelope binning approaches is due to the readily available model diagnostics for random forests, specifically feature importance and partial dependence analysis.

The climatically predicted S_R values from a training dataset consisting of an even mixture of climatically optimal and geologically limited S_R will predict the mean of all pixels present in that climate configuration rather than the climatically optimal one. As a result, the extent to which a low S_R for a given climate is indeed low is underestimated, and therefore absolute deviations between observed and predicted S_R should be interpreted as conservative (minimum) estimates.

A concern with any model is overfitting: if all pixels situated within a certain climate configuration identified by the model are geologically rather than climatically limited, the model will not identify them as having lower than climatically expected S_R because no other other pixels with higher S_R for that climate configuration exist. This weakness is unavoidable with both the random forest approach as well as other empirical climatic envelope binning approaches, but can be overcome to some extent by limiting the decision tree depths (i.e., limiting fit) by enforcing a minimum leaf sample size. Specifying tree depth hyperparameters to limit model fitting comes at the potential cost of absolute model accuracy. However, identification of within-climate configuration variability rather than the best predictive accuracy is the overarching goal in this study. Sensitivity explorations indicated that changing tree depth hyperparameters resulted in varying magnitudes of absolute predicted versus observed S_R differences, nevertheless, the spatial patterns highlighted below were robust.

2.3.4 Geologic layers

We compared the output of the random forest model to existing geologic maps. For statewide analyses, we used the 1:750,000 scale digitized Geologic Map of California (Jennings et al., 2010) to interpret patterns in climatically predicted vs. observed S_R. The map was rasterized to 1 km pixel to match the S_R dataset resolution. We additionally used a 1:65,000 scale geologic map (Huber, 1968) to explore km-scale S_R anomaly patterns across granitic plutons in the Sierra Nevada.

3 Results

Our primary findings are that i) while in general climate can predict S_R with reasonable accuracy, there is substantial unexplained variance; ii) regions where observed S_R tends to be lower than climatically predicted are in many cases spatially bounded by geologic contacts, indicative of a bottom-up geologic control on S_R, and iii) these regions of apparent geologic-controlled S_R are not confined to a particular rock type: diverse lithologies—and hypothesized causal mechanisms—are capable of limiting S_R.

3.1 Observed S_R

Over much of the state, S_R falls between 300 and 600 mm (Figure 2e). The largest observed S_R values (yellow areas in Figure 2e) are found along the western flank of the southern Sierra Nevada and the Transverse Ranges, which also have high interannual variability of precipitation (CV_P, Figure 2c) and moderately high energy delivery (PET, Figure 2a). Very low S_R (purple areas in Figure 2e) is observed in the far north-east (Modoc

Plateau), higher elevation regions in the Sierra Nevada, and parts of the foothills surrounding the Sacramento and Central Valley and the Tulare Basin (the large N-S trending region in white in Figure 2 that was masked from analysis primarily due to large agricultural operations and irrigation).

3.2 Climatically predicted S_R

The random forest model driven by the static climate variables predicts S_R with a root mean square error (RMSE) of 132 mm (the average observed S_R of all pixels is 416 mm) and an R^2 of 0.61. This model was specified to have a minimum of 100 leaf nodes to limit the lumping of particular climate configurations within particular geologic units (see above); hyperparameter tuning estimates indicated that the highest accuracy model would have a minimum of 3 leaf nodes but still have an RMSE of 114 mm. In contrast, a multiple linear regression model including interaction terms (not shown) with the same predictor variables achieves an RMSE of 183 mm, much worse than the random forest. At broad scales, the pattern of predicted S_R using the random forest model (Figure 2f) closely resembles the pattern of observed S_R (Figure 2e).

When the final random forest model is trained with all the available data, analysis of feature importance (Figure 3) indicates that CV_P is the most important predictor of S_R, followed by mean annual P. Thus the random forest model indicates that water supply (its inter-annual variability and average magnitude) are the most important climatic controls on S_R within California, with energy supply (PET) and the intra-annual patterns of water and energy delivery (ASI) being less important.

Partial dependence plots (Figure 4) reveal the marginal effect on predictions of S_R to each climate predictor variable. This analysis indicates that high magnitudes of both P and PET and low magnitudes of CV_P predict low values of S_R. S_R increases monotonically with CV_P, whereas the partial dependence of S_R on P exhibits a humped relationship, with a mesic maximum (Good et al., 2017). There is only a weak negative relationship for ASI. We hypothesize that the physical mechanisms behind these patterns are connected to the impacts of annual magnitude and variability of water delivery. S_R is likely low at low P because there is simply not enough precipitation that arrives prior to dry periods to support much evapotranspiration, limiting the size of the deficit (our measure of S_R) that can grow. S_R is similarly low at high P, but for the opposite reason that locations with high P may have their evapotranspiration limited by energy availability (wetter places tend to have lower potential evapotranspiration in California). S_R may increase with CV_P partly because the denominator in that term is P but also because larger relative inter-annual variability means that plants must rely on more stored water to bridge droughts relative to the typical use for the plant community.

3.3 Regions of climatically under-predicted S_R and underlying geology

While the overall patterns of observed and predicted S_R are similar, the differences reveal where geology may limit plant water availability. Figure 2g shows state-wide areas where the observed S_R is less than the climatically predicted S_R. These pixels are, in many regions, strongly clustered in space and include a large N-S trending swath and other smaller regions of the Northern Coast Ranges, the foothills surrounding the north end of the Sacramento Valley, and large parts of the southern Sierra Nevada. While less obvious in the full map of California, the anomalies are spatially organized at local scales as well (Figure 5 a, d, g, j).

The clustering could be due to a regional, systematic top-down disturbance (e.g., fire, logging, or other unaccounted for land-use) or unaccounted-for climate variable in the model. However, comparison of these regions with geologic mapping indicates instead that substrate is playing the primary role in these spatial patterns.

Figure 5 zooms in on four example regions (one for each row) where S_R anomalies roughly coincide in space with mapped geologic units. The left column of Figure 5 shows how pixels with lower than climatically expected S_R (in dark red) tend to be clustered rather than randomly distributed across the landscape, with clusters aligning reasonably well with outlines of geologic formations. The middle column highlights the particular mapped geologic unit whose extent includes areas of anomalous S_R. The right column shows the same mapped geologic unit's outline superimposed on satellite imagery, and Figure 6 more clearly shows these regions to be less forested than their immediate surroundings. The four highlighted regions have distinct rock types (from top to bottom in Figure 5 and clockwise in Figure 6: melange, volcanic, ultramafic, and granitic). The hypothesized mechanisms for geologic control exerted by each of these rock types is explored in the Discussion below.

In Figure 7, we highlight expansive mapped geologic units (more than 1,000 km^2 areal coverage) where the median of the observed minus predicted S_R is less than -20 mm (i.e., geologic units where the observed S_R tends to be substantially less than the climatically predicted S_R across the state of California). We note that i) these substrates span diverse lithologies (including sedimentary, metamorphic and igneous), and that ii) in some cases, the same units identified visually in the regional case-studies (Figure 5) also exhibit anomalously low S_R at the state-wide scale. Overall, 41% of the study area, or approximately 80,000 km^2 had an observed S_R less than -20 mm than the climatically predicted S_R, indicating that roughly a fifth of California's land area may experience geologically limited storage capacity. It is worth noting that Figure 7 identifies young geologic substrates (Quaternary age) as particularly subject to lower than climatically expected storage capacity. This may be due to a variety of mechanisms, including limited time for nutrients to be fixed or mobilized (Chadwick et al., 1999) or water-retaining clay minerals to form (Jefferson et al., 2010).

4 Discussion

To evaluate where geologic substrates may limit biomass or plant productivity and thus water vapor fluxes to the atmosphere, we identify locations where the observed apparent root-zone water storage capacity (S_R) is smaller than expected relative to other locations with similar climate. Similar to empirical ecological approaches that relate plant productivity or biome characteristics to climate, this empirical identification procedure does not determine the mechanisms underlying the lower-than-expected S_R, which could be associated with disturbance, land-use, or herbivory dynamics. The spatial congruence of many of these locations with geologic boundaries, as opposed to e.g. fire or land use boundaries, provides strong evidence for geologic limitations to plant water availability.

4.1 Process-based mechanisms of geologic limitation of S_R

Figure 1 synthesizes previously proposed mechanisms for geologically limited S_R. Two of these mechanisms are hydrologic mechanisms that limit plant-water availability directly (water storage limitation and water excess) whereas the other two mechanisms indirectly limit S_R via chemical processes that limit plant growth (nutrient limitation and toxicity). We stress that these drivers are not necessarily independent: for example, low nutrient availability could limit plants which in turn limits porosity production in the subsurface. Below, we draw on insights from previous field studies to illustrate how these mechanisms operate, using examples revealed by our mapping as illustrative case studies.

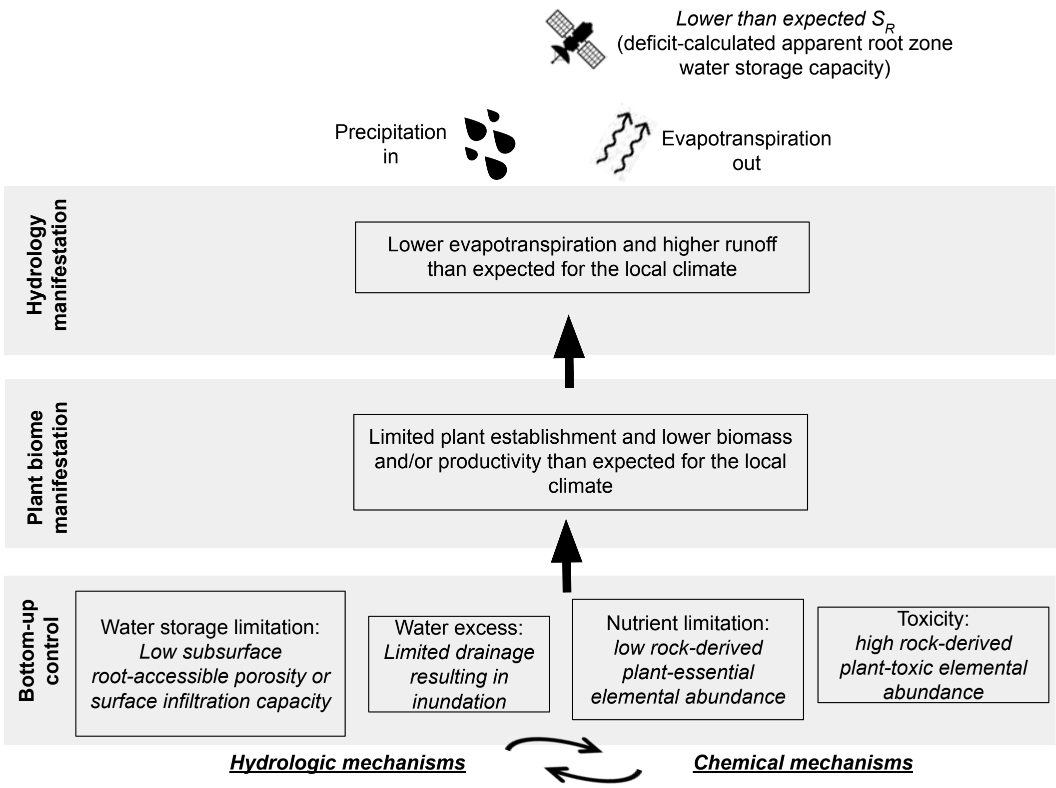

Figure 1. Conceptual diagram illustrating hypothesized geologically mediated controls on apparent root-zone water storage capacity, S_R (lowest row) and corresponding plant biome and hydrologic manifestations. Curved arrows indicate that the geologic controls are not mutually exclusive and may be subject to feedback mechanisms.

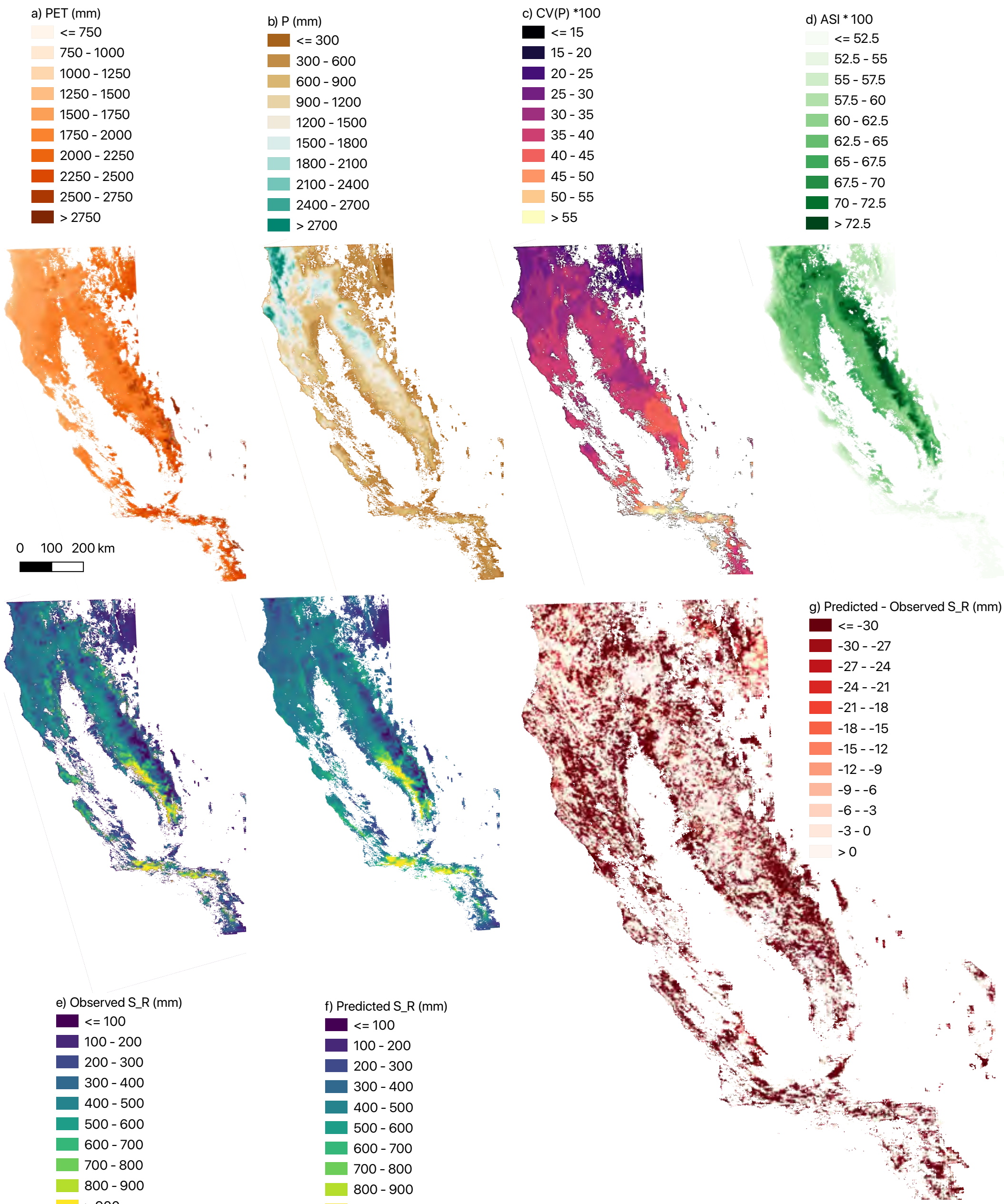

Figure 2. California-wide maps of climatic predictors of S_R (top row) and observed, predicted, and difference between predicted and observed S_R (bottom row). Masked (white) areas are locations where S_R calculation criteria are not met (see Methods).

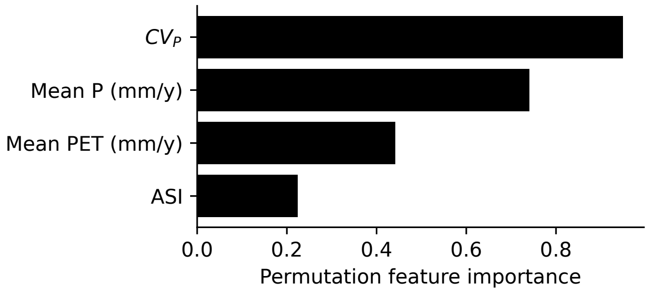

Figure 3. Permutation feature importance of the random forest climate predictors of S_R: higher feature importance indicates that a climate predictor is an important predictor of S_R (inferred by quantifying how much worse the model performs when that variable is randomly shuffled).

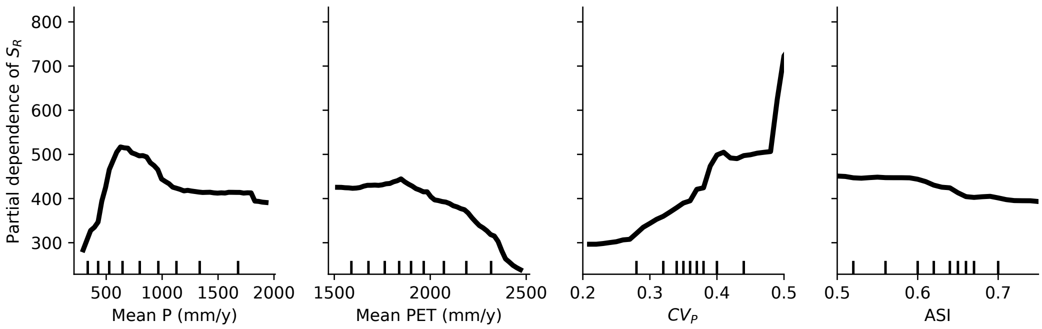

Figure 4. Partial dependence plots show how variation in individual climatic predictor features (x-axes) on average impacts the predicted target variable (S_R, y-axis) when the other climate predictors are controlled for. Vertical lines above x-axes denote decile breaks for the distribution of each climate predictor variable.

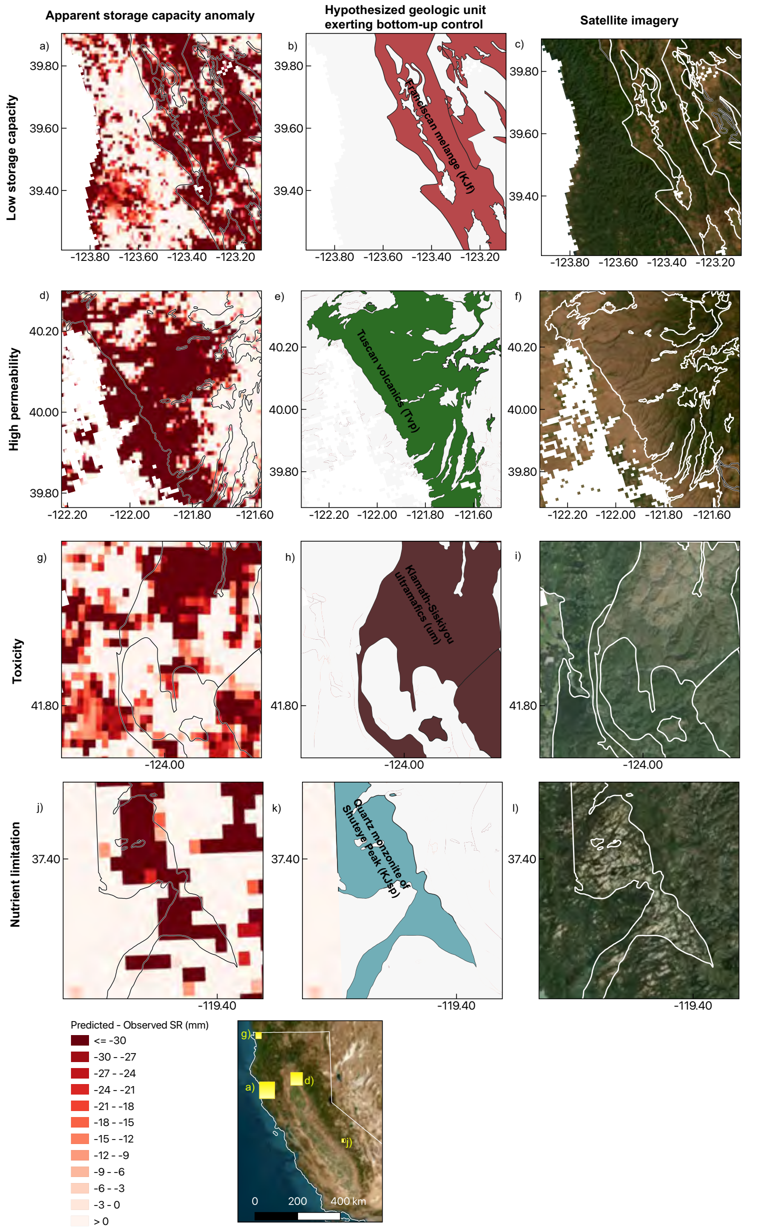

Figure 5. Four regional-scale case studies of apparent geologic control on S_R (one per row). The hypothesized mechanism responsible for anomalously low S_R for the local climate (i.e., red shading in the left column) is identified with the labels at left. The middle column highlights the geologic unit whose spatial extent tends to coincide with a region of anomalous S_R. In the top three rows, the geology mapping comes from the state-wide compilation (Jennings et al., 2010), and in the bottom row from a smaller quadrangle (Huber, 1968). Satellite imagery (from ESRI) in the right column reveals that the low S_R areas also tend to have lower canopy cover than their immediate surroundings. See Discussion for synthesis of prior field studies that support the hypothesized geologic limitation mechanism.

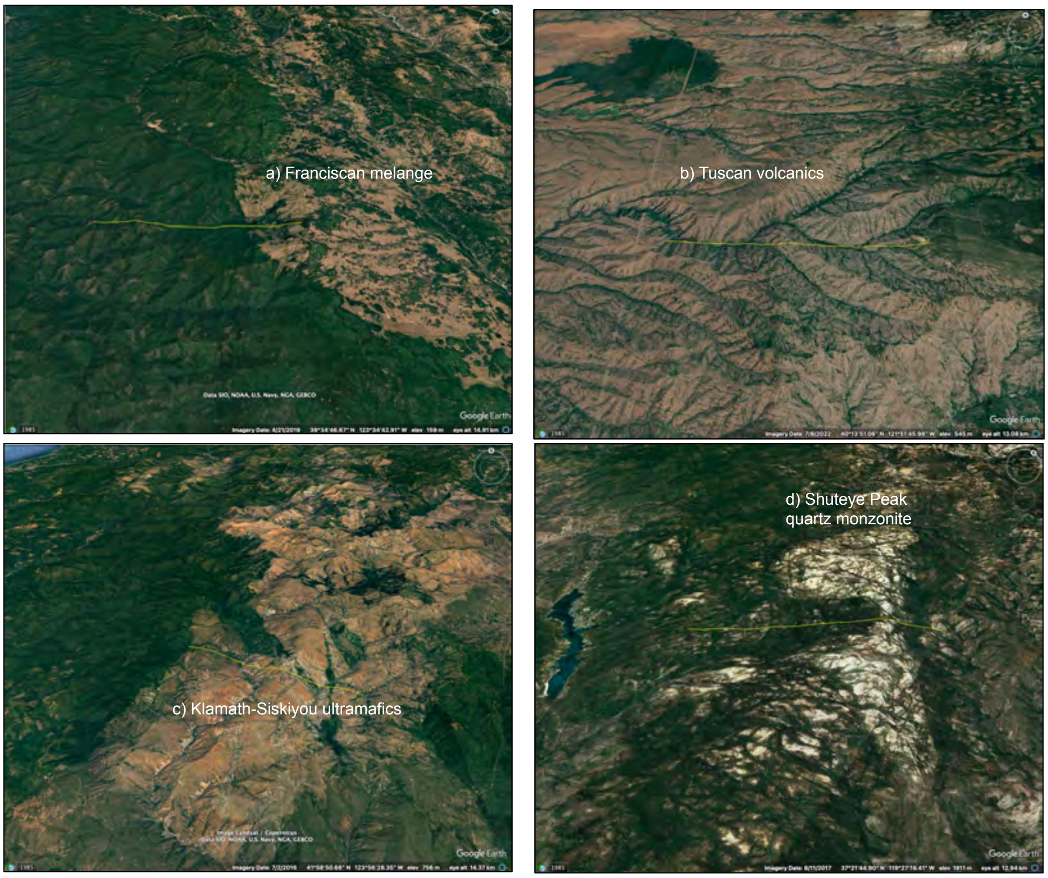

Figure 6. Google Earth imagery with topography of the four case studies highlighted in Figure 5, revealing some of the striking vegetation contrasts over short spatial scales within similar climates that are hypothesized to arise due to geologic controls. The ecotones separating plant communities in these images generally coincide with geologic contacts. In each image, the yellow line is a 10 km scale bar, and the latitude and longitude listed at the lower right of the image is from the center of the scale bar.

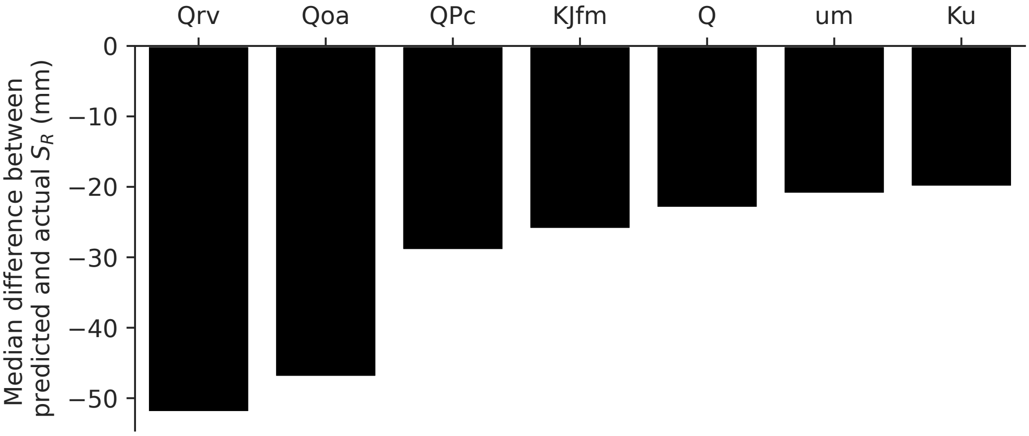

Figure 7. Geologic units with large represented areas ($> 100\ km^2$) that appear to limit root zone storage capacity (i.e., have substantially lower than climatically expected S_R). Key (adapted from Jennings et al. (2010)): Qrv: Volcanic rocks (Holocene) - Recent (Holocene) volcanic flow rocks; minor pyroclastic deposits. Qoa: Marine and nonmarine (continental) sedimentary rocks (Pleistocene) - Older alluvium, lake, playa, and terrace deposits. QPc: Nonmarine (continental) sedimentary rocks (Pleistocene-Holocene) - Pliocene and/or Pleistocene sandstone, shale, and gravel deposits; mostly loosely consolidated. KJfm: Marine sedimentary and metasedimentary rocks (Cretaceous-Jurassic) - Melange of fragmented and sheared Franciscan Complex rocks. Q: Marine and nonmarine (continental) sedimentary rocks (Pleistocene-Holocene) - Alluvium, lake, playa, and terrace deposits; unconsolidated and semi-consolidated. Mostly nonmarine, but includes marine deposits near the coast. um: Plutonic rocks (Mesozoic) - Ultramafic rocks, mostly serpentine. Minor peridotite, gabbro, and diabase; chiefly Mesozoic. Ku: Marine sedimentary and metasedimentary rocks (Upper Cretaceous) - Upper Cretaceous sandstone, shale, and conglomerate.

4.1.1 Water limitation and excess

In both soil and weathered bedrock, connected porosity enables water storage and flow, thereby regulating water status in the root zone (Klos et al., 2018). In upland environments, the weathered bedrock layer is variably thick and typically underlies a physically mobile regolith (soil, in the geomorphological sense) (Rempe & Dietrich, 2014). Weathered bedrock forms as chemical and physical weathering fronts propagate downwards into fresh bedrock as it is nears Earth's surface (Riebe et al., 2017). Under similar climate, spatial gradients in tectonics and lithology can result in variations in weathering extent and thus water storage and flow properties. These variations can result in either limited or excess water, and in some scenarios, both at the same location at different times of year.

For the first case study, we highlight the Central Belt melange of the Franciscan Formation that runs roughly parallel to the coast in the Northern California Coast Ranges (first row in Figure 5). In a region where the local climate can support some of the tallest trees on the planet, the melange is surprisingly sparsely vegetated; instead of the dense forest found immediately to the west, the melange is characterized by an open savanna of deciduous Oregon White oak (*Quercus garryana*) and an herbaceous groundcover (Hahm et al., 2017).

Deep drilling and multiple years of intensive hillslope-scale ecohydrologic field monitoring have resulted in the interpretation that this lower than climatically expected vegetation community arises due the shallow (only 1 to 2 m) propagation of weathering into the fresh melange bedrock (Hahm et al., 2019), which consequently results in limited water storage capacity (about 1/10th of the typical wet season precipitation; in contrast, the Coastal Belt immediately to the west has 20 to 30 m deep weathering fronts and three times greater seasonal water storage (Dralle et al., 2018), with a dense evergreen forest (Figure 6a)). Storage of water from the wet season in the subsurface is critical for plant water supply in the summer dry season in this rain-dominated Mediterranean climate (Hahm, Dralle, et al., 2022). Our mapping in this study extends the insights from hillslope- and catchment-scale field observations and indicates that the melange rock type is associated with lower than climatically expected S_R across the state (the melange is denoted as KJfm in Figure 7).

The low storage capacity of the melange results in both water limitation—in the dry season, when oak pre-dawn water potentials drop below -3 MPa—and water excess, in the wet season, when the subsurface completely saturates repeatedly in storms—resulting in anoxic conditions around flooded roots (Hahm et al., 2018, 2020). The role of excess water as a control on vegetation has also been explored by Sousa et al. (2022); Roebroek et al. (2020); Zipper et al. (2015). The melange presents the interesting situation of rhizosphere water limitation even when a perennially saturated zone is relatively near the surface: in the summer the vadose zone is just a few meters deep, and although the fresh melange beneath is perennially saturated, its extremely low hydraulic conductivity and anoxic conditions apparently prevent root water uptake (Hahm et al., 2020).

In contrast to the scenario where low permeability, perennially saturated fresh bedrock is near the surface, some landscapes can instead have a high conductivity, high porosity substrate that allows infiltrating precipitation to rapidly transit the root zone vertically, draining to deeper aquifers. This form of low vadose zone storage capacity can also lead to water limitation and a lower than climatically expected plant community. These conditions have been documented in karstic terrain in China (H. Liu et al., 2021; Jiang et al., 2020).

We posit that a similar phenomenon may also be possible in highly permeable volcanic bedrock. As a second case study, we highlight a community with low biomass—and low S_R—for the local climate in the Lassen foothills at the north-western end of the

Sacramento Valley (second row in Figure 5). Here, a Pliocene aged volcanic substrate (the Tuscan Formation (Lydon, 1967)) is inhabited by an open oak savanna with abundant rocky outcrops. Both the geomorphology (characterized by buttes) and woody vegetation community (including Interior Live (*Quercus wislizeni* and Blue (*Quercus douglasii*) oaks) are strongly organized along outcrops of particular subhorizontally bedded volcanic deposits (lahars containing tuffs and breccias), as seen in Figure 6b. Based on these bedrock structure and vegetation observations, along with records of high surface infiltration rates and conductivity within permeable beds (Butte County Department of Water and Resource Conservation, 2013), we interpret that in this landscape infiltrating precipitation rapidly transits certain high permeability volcanic beds that comprise the majority of the Formation volumetrically, without significant moisture retention. (Relatively young volcanic landscapes in the Cascades can have relatively little water storage capacity in the near surface and high conductivity (Jefferson et al., 2010; Tague & Grant, 2004) volcanic landscapes). Woody vegetation is minimal on these volcanic beds, but is found along roughly elevation-contour parallel bands where lower conductivity or higher storage capacity beds outcrop at the surface, as vegetation there may experience enhanced water availability from lateral flow or greater retention of infiltrating precipitation.

4.1.2 Toxicity

Toxic concentrations of elements can be released via chemical weathering of underlying bedrock, inhibiting plant growth. Classic examples are associated with ultramafic substrates, and in California there are well-studied examples of high-biodiversity, low-biomass endemic plant communities inhabiting serpentines (Kruckeberg, 1992, 1985; Harrison et al., 2004). In these environments, plants struggle in the presence of exposure to high ratios of Mg:Ca and high Ni (Kruckeberg, 1992).

Consistent with previous observations of low plant biomass on ultramafic substrates, we found that ultramafic areas across the study area tend to have lower than climatically expected S_R (denoted *um* in Figure 7). As a case study, we highlight the dramatic example of a large ultramafic body in the Klamath-Siskiyou region of north-western California—one of the largest in North America (third row in Figure 5 and Figure 6c). This region can climatically support dense evergreen forests, yet the vegetation situated on the serpentine substrate is commonly stunted or altogether absent (Alexander et al., 2007), with scattered individuals of pine, fir and cedar. The inhibited plant growth reduces evapotranspiration, in turn limiting water storage deficits and *apparent* root-zone water storage capacity, as illustrated conceptually in Figure 1. We emphasize that there may in fact be ample water storage capacity, but the stunted plants growing on toxic substrates do not access it, and it is therefore mapped as lower than climatically expected S_R.

4.1.3 Nutrient limitation

Low concentrations of plant-essential nutrients in parent material, low erosion rate and/or high leaching may all contribute to nutrient limitation, stunted vegetation, and lower than expected S_R. In California, nutrient limitation has been associated with ultramafic substrates (see *Toxicity* above), as well as leucogranitic plutons in the Sierra Nevada, where phosphorus concentrations in parent bedrock can be an order of magnitude lower than more mafic adjacent plutons (Hahm et al., 2014). The bottom row of Figure 5 illustrates one such pluton, the Quartz Monzonite of Shuteye Peak, which has low woody plant cover (sparse Jeffrey Pine (*Pinus jeffreyi*)) and large expanses of exposed granitic bedrock, in contrast to nearby granodioritic plutons experiencing a similar climate which are occupied by high biomass evergreen forests, including the charismatic Giant Sequoia (*Sequoiadendron giganteum*); Figure 6d. Ecotones separating the plant communities closely align with mapped intrusive contacts (Huber, 1968; Hahm et al., 2014). Neither Shuteye Peak nor the nearby Bald Mountain were glaciated in the

Pleistocene, and their sparse soil cover has been attributed to nutrient limitation that inhibits root growth which consequently inhibits soil retention (Hahm et al., 2014). This has been hypothesized to result in a feedback cycle that further inhibits weathering and porosity production in the subsurface, which in turn also limits the water storage capacity for trees and their growth (Callahan et al., 2022). Thus, the nutrient and water limitation geologic controls on plant abundance, water use, and ultimately S_R are potentially closely linked via feedback cycles, defining an exciting research frontier.

4.2 Implications for climate change driven plant biome migration and the use of S_R in models

Bioclimatic modeling approaches provide a first approximation to the availability of plant habitat (Pearson & Dawson, 2003). It has long been argued, however, that physiographic, edaphic, and geophysical factors—in addition to climate—should be taken into consideration when predicting and managing for climate change induced species migration (Theobald et al., 2015; Anderson & Ferree, 2010; Hulshof & Spasojevic, 2020; Davis et al., 2018; Butler et al., 2007; Macias-Fauria & Johnson, 2013), a sentiment well captured by Kruckeberg (2013): "given a regional climatic framework, much of the plant species diversity and discontinuity in the region is governed by variations in soil chemistry, and thus by specific variations in the mineralogy of rock substrates." Our work builds on these insights by enabling a direct quantification of the impact of geology over large spatial scales using recently made available, spatially distributed estimates of S_R and a simple, climate-driven machine learning model.

S_R is a key parameter across hydrology, vegetation, and climate models (Seneviratne et al., 2013), because of its large impact on terrestrial water partitioning, plant-water availability and associated carbon uptake, and the associated impacts of latent heat flux and vegetation greenness on the climate. Although previous studies have used both climate and soils databases to establish edaphoclimatic envelopes for modeling vegetation distribution (de Castro Oliveira et al., 2021), there is a growing consensus that traditionally used static soils database derived estimates of S_R are inadequate (Stocker et al., 2023). This is due to the mounting evidence of widespread plant-water uptake from bedrock whose water storage properties are not traditionally included within soils databases (McCormick et al., 2021; Stocker et al., 2023) and because temporally changing vegetation communities can result in shifting magnitudes of S_R at a single location (L. Zhang et al., 2001; Li et al., 2019; Nijzink et al., 2016; Hrachowitz et al., 2021). Our approach offers a path forward for *empirically identifying* geologic limitations on S_R, but we do not see a clear way to *predict* such limitations a priori at large spatial scales at the moment, particularly when they arise due to hydrologic mechanisms (Figure 1). This is due to complicated feedbacks among the various processes and our current inability to directly observe weathering extent and water storage and flow properties at large spatial scales.

4.3 Limitations and future work

The distinction between top-down (climate) versus bottom-up (geologic) drivers of S_R becomes murky over longer time scales. This is partly because landscapes inherit paleoclimate weathering legacies: for example, climate may result in glaciation, which can strip away soil and weathered bedrock, resulting in a proximate bottom-up control on S_R that is facilitated ultimately by a long-term climate history. Climate drivers are also filtered by the subsurface to determine groundwater dynamics, which can strongly impact plant community distribution over individual hillslope lengthscales (Koirala et al., 2017; Roebroek et al., 2020; Fan et al., 2017). Climate also impacts hillslope diffusive and advective erosive processes, which may impact seedling establishment (Toloui-Semnani & Johnson, 2019), the thickness of the weathered bedrock zone and the sizes of colluvial wedges (and potential storage space for water (Ding et al., 2018; Milodowski et al., 2015; Rempe & Dietrich, 2014)) and the spacing of ridges and valleys (Perron et al., 2009). It has also been argued that vegetation 'coevolves' with the subsurface in such a way to produce a particular water storage reservoir: in this view, soils are largely biotic constructs (van Breemen, 1993). The approach outlined in this study is not capable of teasing apart the longer-term connections between top-down and bottom-up drivers of S_R—instead, it takes the current climate at face value and asks whether the empirically observed S_R is lower in some places relative to others with the same climate. While this works in many locations (e.g., the case studies explored above), this empirical approach is incapable of detecting a bottom-up limitation on S_R if all locations for a particular climate are similarly limited by a geologically mediated factor.

An additional complication in identifying bottom-up limitations of S_R can arise in locations with significant inter-pixel lateral groundwater subsidies to vegetation (Roebroek et al., 2020; Fan et al., 2017). In this scenario, a larger than climatically expected S_R may be detected because evapotranspiration is sustained by groundwater flow from elsewhere, which could result in large calculated water storage deficits. We expect this process to be most common at the scale of individual hillslopes, where water that infiltrates near local topographic highs may flow laterally downslope toward local channels. Because the pixel sizes we consider are large relative to local hillslope lengthscales, however, this effect should be minimized in our estimation procedure.

5 Conclusions

We employed a simple machine learning approach to quantify the difference between climatically expected and observed apparent root-zone water storage capacity (S_R). By comparing the resulting patterns with geologic maps, we found strong spatial correspondence between particular substrates and regions of lower than climatically expected S_R. These patterns are indicative of bottom-up controls on the size of the root zone. Our mapping approach is not capable of identifying the mechanisms by which geology limits S_R. However, the patterns we observed are consistent with mechanisms identified in previous field studies, which highlight the role of water availability (excess and limitation), nutrient supply, and toxicity. Although our analysis is not exhaustive, the approach presented here enables extension of hillslope-scale field inferences to much larger areas, and, importantly, does not rely on traditionally used soil water storage capacity databases, which are generally too shallow to capture relevant plant water dynamics in seasonally dry climates. Furthermore, our findings indicate that climate patterns alone can be insufficient predictors of root zone water storage capacity. The subsurface matters, and should be incorporated into earth system models and ecosystem migration management plans in the context of climate change.

6 Open Research

All data sets used in this research were previously published (see references in Methods for details). The Python notebooks used to query Google Earth Engine, aggregate data, and perform the random forest modeling and other data analyses are available on Hydroshare: `https://www.hydroshare.org/resource/be4e3be9e18144908bd4a7baa75a9a4e/` (Hahm, 2023).

Acknowledgments

Funding support from Simon Fraser University, the Canadian Foundation for Innovation, and a Natural Sciences and Engineering Research Council of Canada Discovery Grant.

References

Alexander, E. B., Coleman, R. G., Keeler-Wolfe, T., & Harrison, S. P. (2007). Ser-

pentine plant life of western north america. In *Serpentine geoecology of western north america.* Oxford University Press.
Anderson, M. G., & Ferree, C. E. (2010). Conserving the stage: climate change and the geophysical underpinnings of species diversity. *PloS one*, *5*(7), e11554.
Bouaziz, L. J., Aalbers, E. E., Weerts, A. H., Hegnauer, M., Buiteveld, H., Lammersen, R., ... Hrachowitz, M. (2022). Ecosystem adaptation to climate change: the sensitivity of hydrological predictions to time-dynamic model parameters. *Hydrology and Earth System Sciences*, *26*(5), 1295–1318.
Butler, D. R., Malanson, G. P., Walsh, S. J., & Fagre, D. B. (2007). Influences of geomorphology and geology on alpine treeline in the american west—more important than climatic influences? *Physical Geography*, *28*(5), 434–450.
Butte County Department of Water and Resource Conservation. (2013, Feb). *Lower tuscan aquifer monitoring, recharge, and data management project.* Retrieved from `https://www.buttecounty.net/Portals/26/Tuscan/LTAFinalReport.pdf`
Callahan, R. P., Riebe, C. S., Sklar, L. S., Pasquet, S., Ferrier, K. L., Hahm, N. J., W aJnd Taylor, ... others (2022). Forest vulnerability to drought controlled by bedrock composition. *Nature Geoscience*, *15*(9), 714–719.
Chadwick, O. A., Derry, L. A., Vitousek, P. M., Huebert, B. J., & Hedin, L. O. (1999). Changing sources of nutrients during four million years of ecosystem development. *Nature*, *397*(6719), 491–497.
Cheng, S., Cheng, L., Qin, S., Zhang, L., Liu, P., Liu, L., ... Wang, Q. (2022). Improved understanding of how catchment properties control hydrological partitioning through machine learning. *Water Resources Research*, *58*(4), e2021WR031412.
Daly, C., Smith, J. I., & Olson, K. V. (2015). Mapping atmospheric moisture climatologies across the conterminous united states. *PloS one*, *10*(10), e0141140.
Davis, E. L., Hager, H. A., & Gedalof, Z. (2018). Soil properties as constraints to seedling regeneration beyond alpine treelines in the canadian rocky mountains. *Arctic, Antarctic, and Alpine Research*, *50*(1), e1415625.
Dawson, T. E., Hahm, W., & Crutchfield-Peters, K. (2020). Digging deeper: what the critical zone perspective adds to the study of plant ecophysiology. *New Phytologist*, *226*(3), 666–671.
de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C., & Savenije, H. H. (2016). Influence of soil and climate on root zone storage capacity. *Water Resources Research*, *52*(3), 2009–2024.
de Boer-Euser, T., Palalane, J., Savenije, H., & Juízo, D. (2019). How climate variations are reflected in root zone storage capacities. *Physics and Chemistry of the Earth, Parts A/B/C*, *112*, 83–90.
de Castro Oliveira, G., Arruda, D. M., Fernandes Filho, E. I., Veloso, G. V., Francelino, M. R., & Schaefer, C. E. G. R. (2021). Soil predictors are crucial for modelling vegetation distribution and its responses to climate change. *Science of The Total Environment*, *780*, 146680.
Ding, J., Johnson, E. A., & Martin, Y. E. (2018). Linking soil moisture variation and abundance of plants to geomorphic processes: A generalized model for erosion-uplifting landscapes. *Journal of Geophysical Research: Biogeosciences*, *123*(3), 960–975.
Dralle, D. N., Hahm, W., Chadwick, K. D., McCormick, E., & Rempe, D. M. (2021). Accounting for snow in the estimation of root zone water storage capacity from precipitation and evapotranspiration fluxes. *Hydrology and Earth System Sciences*, *25*(5), 2861–2867.
Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., & Dietrich, W. E. (2018). Quantification of the seasonal hillslope water storage that does not drive streamflow. *Hydrological processes*, *32*(13), 1978–1992.
Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C.

(2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, *114*(40), 10572–10577.

Feng, X., Thompson, S. E., Woods, R., & Porporato, A. (2019). Quantifying asynchronicity of precipitation and potential evapotranspiration in mediterranean climates. *Geophysical Research Letters*, *46*(24), 14692–14701.

Gao, H., Hrachowitz, M., Schymanski, S., Fenicia, F., Sriwongsitanon, N., & Savenije, H. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, *41*(22), 7916–7923.

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the budyko curve. *Geophysical Research Letters*, *39*(19).

Good, S. P., Moore, G. W., & Miralles, D. G. (2017). A mesic maximum in biological water use demarcates biome sensitivity to aridity shifts. *Nature ecology & evolution*, *1*(12), 1883–1888.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18–27.

Guswa, A. J. (2008). The influence of climate on root depth: A carbon cost-benefit analysis. *Water Resources Research*, *44*(2).

Guswa, A. J. (2010). Effect of plant uptake strategy on the water- optimal root depth. *Water Resources Research*, *46*(9).

Hahm, W. (2023). *Data and code for: "geologic controls on apparent root-zone storage capacity"*. Hydroshare. Retrieved from `https://www.hydroshare.org/resource/be4e3be9e18144908bd4a7baa75a9a4e/`

Hahm, W., Dietrich, W. E., & Dawson, T. E. (2018). Controls on the distribution and resilience of quercus garryana: ecophysiological evidence of oak's water-limitation tolerance. *Ecosphere*, *9*(5), e02218.

Hahm, W., Dralle, D., Lovill, S., Rose, J., Dawson, T., & Dietrich, W. (2017). Exploratory tree survey (2016—eel river critical zone observatory—sagehorn—central belt melange, franciscan complex, northern california coast ranges, usa). hydroshare. *HydroShare https://doi. org/10.4211/hs. 7881821a5c0e4ae3822b96a59f4bf8b6*.

Hahm, W., Dralle, D., Sanders, M., Bryk, A., Fauria, K., Huang, M., et al. (2022). Bedrock vadose zone storage dynamics under extreme drought: Consequences for plant water availability, recharge, and runoff. *Water Resources Research*, *58*, e2021WR031781.

Hahm, W., Lapides, D., Rempe, D., McCormick, E., & Dralle, D. (2022). The age of evapotranspiration: Lower-bound constraints from distributed water fluxes across the continental united states. *Water Resources Research*, *58*(10), e2022WR032961.

Hahm, W., Rempe, D., Dralle, D., Dawson, T., & Dietrich, W. (2020). Oak transpiration drawn from the weathered bedrock vadose zone in the summer dry season. *Water Resources Research*, *56*(11), e2020WR027419.

Hahm, W., Rempe, D. M., Dralle, D. N., Dawson, T. E., Lovill, S. M., Bryk, A. B., ... Dietrich, W. E. (2019). Lithologically controlled subsurface critical zone thickness and water storage capacity determine regional plant community composition. *Water Resources Research*, *55*(4), 3028–3055.

Hahm, W., Riebe, C. S., Lukens, C. E., & Araki, S. (2014). Bedrock composition regulates mountain ecosystems and landscape evolution. *Proceedings of the National Academy of Sciences*, *111*(9), 3338–3343.

Hall, D. K., Riggs, G. A., Foster, J. L., & Kumar, S. V. (2010). Development and evaluation of a cloud-gap-filled modis daily snow-cover product. *Remote sensing of environment*, *114*(3), 496–503.

Harrison, S., Safford, H., & Wakabayashi, J. (2004). Does the age of exposure of ser-

pentine explain variation in endemic plant diversity in california? *International Geology Review*, *46*(3), 235–242.
Holbrook, W. S., Riebe, C. S., Elwaseif, M., L. Hayes, J., Basler-Reeder, K., L. Harry, D., ... W. Hopmans, J. (2014). Geophysical constraints on deep weathering and water storage potential in the southern sierra critical zone observatory. *Earth Surface Processes and Landforms*, *39*(3), 366–380.
Hrachowitz, M., Stockinger, M., Coenders-Gerrits, M., van der Ent, R., Bogena, H., Lücke, A., & Stumpp, C. (2021). Reduction of vegetation-accessible water storage capacity after deforestation affects catchment travel time distributions and increases young water fractions in a headwater catchment. *Hydrology and earth system sciences*, *25*(9), 4887–4915.
Huber, N. K. (1968). *Geologic map of the shuteye peak quadrangle, sierra nevada, california.* Retrieved from https://catalog.data.gov/dataset/digital-geologic-gis-map-of-the-shuteye-peak-15-quadrangle-california-nps-grd-gri-yose-shu
Hulshof, C. M., & Spasojevic, M. J. (2020). The edaphic control of plant diversity. *Global Ecology and Biogeography*, *29*(10), 1634–1650.
Jefferson, A., Grant, G., Lewis, S., & Lancaster, S. (2010). Coevolution of hydrology and topography on a basalt landscape in the oregon cascade range, usa. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *35*(7), 803–816.
Jennings, C. W., Gutierrez, C., Bryant, W., Saucedo, G., & Wills, C. (2010, Dec). *Geologic map of california version 2.0.* Retrieved from https://maps.conservation.ca.gov/cgs/#datalist
Jiang, Z., Liu, H., Wang, H., Peng, J., Meersmans, J., Green, S. M., ... Song, Z. (2020). Bedrock geochemistry influences vegetation growth by regulating the regolith water holding capacity. *Nature communications*, *11*(1), 2392.
Keppeler, E. T., & Ziemer, R. R. (1990). Logging effects on streamflow: water yield and summer low flows at caspar creek in northwestern california. *Water resources research*, *26*(7), 1669–1679.
Kleidon, A., & Heimann, M. (1998). A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle. *Global Change Biology*, *4*(3), 275–286.
Klos, P. Z., Goulden, M. L., Riebe, C. S., Tague, C. L., O'Geen, A. T., Flinchum, B. A., ... others (2018). Subsurface plant-accessible water in mountain ecosystems with a mediterranean climate. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1277.
Koirala, S., Jung, M., Reichstein, M., de Graaf, I. E., Camps-Valls, G., Ichii, K., ... others (2017). Global distribution of groundwater-vegetation spatial covariation. *Geophysical Research Letters*, *44*(9), 4134–4142.
Kruckeberg, A. (1985). *California serpentines: flora, vegetation, geology, soils, and management problems.* Univ of California Press.
Kruckeberg, A. (1992). Plant life of western north american ultramafics. *The ecology of areas with serpentinized rocks: a world view*, 31–73.
Kruckeberg, A. (2013). Soil diversity and the distribution of plants, with examples from western north america. *Madroño*, *60*(4), 267–292.
Kuijper, D. P., Te Beest, M., Churski, M., & Cromsigt, J. (2015). Bottom-up and topdown forces shaping wooded ecosystems: lessons from a cross-biome comparison. *Trophic Ecology. Cambridge University Press, Cambridge*, 107–133.
Lapides, D. A., Hahm, W., Rempe, D. M., Whiting, J., & Dralle, D. N. (2022). Causes of missing snowmelt following drought. *Geophysical Research Letters*, *49*(19), e2022GL100505.
Li, H., Si, B., Wu, P., & McDonnell, J. J. (2019). Water mining from the deep critical zone by apple trees growing on loess. *Hydrological Processes*, *33*(2), 320–327.

Liu, H., Dai, J., Xu, C., Peng, J., Wu, X., & Wang, H. (2021). Bedrock-associated belowground and aboveground interactions and their implications for vegetation restoration in the karst critical zone of subtropical southwest china. *Progress in Physical Geography: Earth and Environment*, *45*(1), 7–19.

Liu, M., Zhang, B., & He, X. (2022). Climate rather than vegetation changes dominate changes in effective vegetation available water capacity. *Water Resources Research*, *58*(3), e2021WR030319.

Lydon, P. A. (1967). The origin of the tuscan buttes and the volume of the tuscan formation in northern california. *Short Contributions to California Geology: Special Report*, *91*, 17–26.

Macias-Fauria, M., & Johnson, E. A. (2013). Warming-induced upslope advance of subalpine forest is severely limited by geomorphic processes. *Proceedings of the National Academy of Sciences*, *110*(20), 8117–8122.

McCormick, E. L., Dralle, D. N., Hahm, W., Tune, A. K., Schmidt, L. M., Chadwick, K. D., & Rempe, D. M. (2021). Widespread woody plant use of water stored in bedrock. *Nature*, *597*(7875), 225–229.

McLaughlin, B. C., Blakey, R., Weitz, A. P., Feng, X., Brown, B. J., Ackerly, D. D., ... Thompson, S. E. (2020). Weather underground: Subsurface hydrologic processes mediate tree vulnerability to extreme climatic drought. *Global change biology*, *26*(5), 3091–3107.

Milly, P. (1994). Climate, interseasonal storage of soil water, and the annual water balance. *Advances in Water Resources*, *17*(1-2), 19–24.

Milodowski, D. T., Mudd, S. M., & Mitchard, E. T. (2015). Erosion rates as a potential bottom-up control of forest structural characteristics in the sierra nevada mountains. *Ecology*, *96*(1), 31–38.

Morford, S. L., Houlton, B. Z., & Dahlgren, R. A. (2011). Increased forest ecosystem carbon and nitrogen storage from nitrogen rich bedrock. *Nature*, *477*(7362), 78–81.

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., ... others (2016). The evolution of root-zone moisture capacities after deforestation: a step towards hydrological predictions under change? *Hydrology and Earth System Sciences*, *20*(12), 4775–4799.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global ecology and biogeography*, *12*(5), 361–371.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perron, J. T., Kirchner, J. W., & Dietrich, W. E. (2009). Formation of evenly spaced ridges and valleys. *Nature*, *460*(7254), 502–505.

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *The American Naturalist*, *164*(5), 625–632.

Rempe, D. M., & Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy of Sciences*, *111*(18), 6576–6581.

Riebe, C. S., Hahm, W., & Brantley, S. L. (2017). Controls on deep critical zone architecture: A historical review and four testable hypotheses. *Earth Surface Processes and Landforms*, *42*(1), 128–156.

Roebroek, C. T., Melsen, L. A., Hoek van Dijke, A. J., Fan, Y., & Teuling, A. J. (2020). Global distribution of hydrologic controls on forest growth. *Hydrology and Earth System Sciences*, *24*(9), 4625–4639.

Running, S., Mu, Q., & Zhao, M. (2017). Mod16a2 modis/terra net evapotranspiration 8-day l4 global 500m sin grid v006. *NASA EOSDIS Land Processes DAAC*, *6*.

Schenk, H. J. (2008). The shallowest possible water extraction profile: A null model for global root distributionsall rights reserved. no part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. *Vadose Zone Journal*, *7*(3), 1119–1124.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., & Hutley, L. B. (2008). An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and earth system sciences*, *12*(3), 913–932.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., ... others (2013). Impact of soil moisture-climate feedbacks on cmip5 projections: First results from the glace-cmip5 experiment. *Geophysical Research Letters*, *40*(19), 5212–5217.

Singh, C., Wang-Erlandsson, L., Fetzer, I., Rockström, J., & Van Der Ent, R. (2020). Rootzone storage capacity reveals drought coping strategies along rainforest-savanna transitions. *Environmental Research Letters*, *15*(12), 124021.

Sousa, T. R., Schietti, J., Ribeiro, I. O., Emílio, T., Fernández, R. H., ter Steege, H., ... others (2022). Water table depth modulates productivity and biomass across amazonian forests. *Global Ecology and Biogeography*.

Speich, M. J., Lischke, H., & Zappa, M. (2018). Testing an optimality-based model of rooting zone water storage capacity in temperate forests. *Hydrology and Earth System Sciences*, *22*(7), 4097–4124.

Stocker, B. D., Tumber-Dávila, S. J., Konings, A. G., Anderson, M. C., Hain, C., & Jackson, R. B. (2023). Global patterns of water storage in the rooting zones of vegetation. *Nature Geoscience*, 1–7.

Tague, C., & Grant, G. E. (2004). A geological framework for interpreting the low-flow regimes of cascade streams, willamette river basin, oregon. *Water Resources Research*, *40*(4).

Theobald, D. M., Harrison-Atlas, D., Monahan, W. B., & Albano, C. M. (2015). Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PloS one*, *10*(12), e0143619.

Toloui-Semnani, M., & Johnson, E. A. (2019). Processes that cause natural unforested areas in canadian rockies below temperature treeline. *Geomorphology*, *347*, 106857.

Trautmann, T., Koirala, S., Carvalhais, N., Güntner, A., & Jung, M. (2022). The importance of vegetation in understanding terrestrial water storage variations. *Hydrology and Earth System Sciences*, *26*(4), 1089–1109.

van Breemen, N. (1993). Soils as biotic constructs favouring net primary productivity. *Geoderma*, *57*(3), 183–211.

van Oorschot, F., van der Ent, R. J., Hrachowitz, M., & Alessandri, A. (2021). Climate-controlled root zone parameters show potential to improve water flux simulations by land surface models. *Earth System Dynamics*, *12*(2), 725–743.

Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I., ... Savenije, H. H. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrology and Earth System Sciences*, *20*(4), 1459–1481.

Witty, J. H., Graham, R. C., Hubbert, K. R., Doolittle, J. A., & Wald, J. A. (2003). Contributions of water supply from the weathered bedrock zone to forest soil quality. *Geoderma*, *114*(3-4), 389–400.

Yang, Y., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, *52*(10), 8260–8276.

Zhang, C., Wang, Y., Jia, X., An, Z., et al. (2020). Variations in capacity and storage of plant-available water in deep profiles along a revegetation and precipitation gradient. *Journal of Hydrology*, *581*, 124401.

Zhang, L., Dawes, W., & Walker, G. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water resources research*, *37*(3), 701–708.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H., McVicar, T. R., Zhang, Q., & Yang, Y. (2019). Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote sensing of environment*, *222*, 165–182.

Zhu, X., Liu, H., He, W., Wu, L., & Liu, F. (2023). Regolith water storage patterns determine vegetation productivity in global karst regions. *Geoderma*, *430*, 116292.

Zipper, S. C., Soylu, M. E., Booth, E. G., & Loheide, S. P. (2015). Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. *Water Resources Research*, *51*(8), 6338–6358.

Geologic controls on apparent root-zone storage capacity

W.J. Hahm[1], D.N. Dralle[2], D.A. Lapides[1,2], R.S. Ehlert[1], D.M. Rempe[3]

[1]Simon Fraser University, Burnaby, BC, Canada

[2]Pacific Southwest Research Station, United States Forest Service, Davis, CA, USA

[3]University of Texas at Austin, Austin, TX, USA

Key Points:

- Regionally extensive areas of low apparent root-zone storage capacity for a particular climate coincide with particular geologic substrates
- Hypothesized geologic controls include water storage capacity limitation, nutrient limitation, and toxicity

Corresponding author: WJ Hahm, `whahm@sfu.ca`

Abstract

The water storage capacity of the root zone determines whether plants survive dry periods and controls the partitioning of precipitation into streamflow and evapotranspiration. It is currently thought that top-down, climatic factors are the primary control on this capacity via their interaction with plant rooting adaptations. However, it remains unclear to what extent bottom-up, geologic factors can provide an additional constraint on storage capacity. Here we use a machine learning approach to identify regions with lower than climatically expected apparent storage capacity. We find that in seasonally dry California these regions overlap with particular geologic substrates. We hypothesize that these patterns reflect diverse mechanisms by which substrate can limit storage capacity, and highlight case studies consistent with limited weathered bedrock extent (melange in the Northern Coast Range), toxicity (ultramafic substrates in the Klamath-Siskiyou region), nutrient limitation (phosphorus-poor plutons in the southern Sierra Nevada), and low porosity capable of retaining water (volcanic formations in the southern Cascades). The observation that at regional scales climate alone does not 'size' the root zone has implications for the parameterization of storage capacity in models of plant dynamics (and the interrelated carbon and water cycles), and also underscores the importance of geology in considerations of climate-change induced biome migration and habitat suitability.

Plain Language Summary

What determines how much water plants can store in their root zone? One school of thought posits that plants 'size' the root-zone capacity to survive a drought of a particular return period. In this scenario, plants extend their roots into the subsurface in response to climate drivers (e.g., precipitation magnitude-frequency and atmospheric water demand). This worldview neglects the potential for geology to restrict root access to water. 'Bottom-up' limitations on storage capacity have been described at individual field sites, but it has been unclear how to identify geologic limitations at large scales. Here, we introduce an approach that quantifies differences between the climatically expected and locally observed apparent storage capacity, and relate these spatial patterns to geologic substrate. Importantly, we quantify apparent storage capacity via a method that includes water below the upper 1.5 m, within weathered bedrock, which is an important water source in seasonally dry climates and is typically excluded from traditional soil texture databases. We find that geology limits storage capacity at regional scales, and synthesize existing field evidence to hypothesize mechanisms of bottom-up control. Our findings have important implications for water-carbon cycle modeling efforts and the prediction of plant biome migration in response to climate change.

1 Introduction

Root-accessible water storage capacity in the subsurface is a key earth system property that regulates the water and carbon cycles (Kleidon & Heimann, 1998). For example, plant transpiration of stored water is a first-order control on Earth's surface energy budget and terrestrial water partitioning (Milly, 1994), setting aquatic ecosystem habitat and water quality and quantity for downstream users. Sufficient storage capacity also enables plants to bridge meteorologic droughts and sustain photosynthesis during extreme dry periods (Porporato et al., 2004; McLaughlin et al., 2020). It has been argued that top-down (climatic) drivers are primarily responsible for determining the large-scale spatiotemporal variability of storage capacity (Nijzink et al., 2016; Guswa, 2008, 2010; M. Liu et al., 2022; van Oorschot et al., 2021; Bouaziz et al., 2022). However, field investigations have revealed that geologic or edaphic factors can exert a primary control at some sites (e.g., Hahm et al., 2019), but it is presently unknown where and why geologic factors eclipse climate factors at landscape scales. This uncertainty challenges earth system and dynamic global vegetation modeling efforts, including prediction of plant biome migration in the context of climate change.

Plant-available water storage capacity is understood to be set by i) the porosity profile, which determines the amount of water that can be held at various water potentials, and ii) the presence of roots to access that porosity (Klos et al., 2018; C. Zhang et al., 2020). Factors related to geology can limit the storage capacity, either directly (by restricting accessible porosity in the near surface (e.g, the presence of low porosity fresh bedrock at a shallow depth) (Hahm et al., 2019) or by being too permeable to store water (H. Liu et al., 2021; Jiang et al., 2020)) or indirectly (by inhibiting plant growth (via nutrient limitation or toxicity) (Hahm et al., 2014; Kruckeberg, 1985; Morford et al., 2011) that in turn inhibits root exploitation of accessible porosity), as depicted in Figure 1. In contrast, top-down (climatic) controls are thought to determine the storage capacity primarily by setting atmospheric water demand and precipitation inputs, including the frequency and duration of dry periods that plants need to endure to survive. Various models that explore optimal plant strategy suggest that plants will invest just enough carbon in root profiles to have sufficient water access to survive dry periods of a particular recurrence interval (Schymanski et al., 2008; Schenk, 2008; Yang et al., 2016; Speich et al., 2018; Guswa, 2008, 2010). This school of thought is encapsulated in the notion that climate 'sizes' the root-zone storage capacity (Gao et al., 2014; de Boer-Euser et al., 2016, 2019; Gentine et al., 2012). Optimal rooting frameworks may neglect the potential for bottom-up factors to limit storage capacity, however, because they implicitly treat the subsurface like an infinite sand box, into which plants may invest as much—or as little—into rooting as is advantageous (e.g., Singh et al., 2020).

A first-order challenge in teasing apart climatic versus geologic controls on storage capacity is quantifying the actual storage capacity accessed by plants. Traditionally, the storage capacity has been parameterized in models through calibration or with the aid of distributed soils datasets, which typically quantify water retention properties through the upper 1 to 1.5 m or to the depth of a restrictive layer. Although widely available and relatively finely resolved, soils datasets have two principle shortcomings: i) they do not capture whether roots are actually present in the soil profile, and ii) they do not extend deep enough into the subsurface to capture porosity profiles in deeper weathered bedrock that commonly underlies soils (Holbrook et al., 2014; Witty et al., 2003; Dawson et al., 2020), where widespread evidence has emerged of root penetration and water uptake (McCormick et al., 2021; Zhu et al., 2023; Stocker et al., 2023). The relative inaccessibility of the deep subsurface challenges quantification of these factors (Stocker et al., 2023).

A recently developed and now widely adopted alternative approach (Wang-Erlandsson et al., 2016; Dralle et al., 2021) constrains storage capacity via tracking of hydrologic fluxes. Precipitation (flux in) and evapotranspiration (flux out) are monitored at a location, and it is reasoned that the root-accessible subsurface water storage capacity must be big enough to explain the largest observed cumulative evapotranspiration in excess of precipitation over a period of record (i.e., the largest observed storage deficit). This approach quantifies an *apparent* root-zone water storage capacity (S_R): i.e., S_R identified from the largest observed deficit is only a lower bound on actual accessible storage capacity (McCormick et al., 2021). For example, it is possible that plants may have had access to—and would have used—more water if dry conditions persisted. In other words, actual root-zone water storage capacity may be larger than S_R, but we do not have the means to directly measure it (although some researchers have attempted to quantify it by fitting yearly maximum deficit values to extreme value distributions (Wang-Erlandsson et al., 2016)). Nevertheless, storage capacity calculated via deficit-style approaches has many theoretical and pragmatic advantages. S_R results in improved hydrological model performance when used as an input parameter (Wang-Erlandsson et al., 2016) and can explain continental-scale patterns in water partitioning (Cheng et al., 2022) and storage dynamics (Trautmann et al., 2022); deficit calculations have also proven essential in the accurate prediction of snowmelt contributions to streamflow following droughts (Lapides et al., 2022). Importantly, deficit-calculated S_R does not require *a priori* assumptions regarding porosity or rooting profiles; distributed hydrologic flux datasets make it feasible to estimate S_R at large spatial scales in cloud-based analysis platforms like Google Earth Engine.

Although distributed estimates of S_R are now available, it has remained challenging to isolate both the spatial patterns and drivers of geologic factors impacting the magnitude of S_R. Here we explore a simple machine learning approach to predict S_R, assuming that climatic controls are the primary drivers of spatial variations in S_R. This exercise reveals locations where the null hypothesis may be rejected (i.e., places where geologic controls may be important) based on deviations between the S_R predicted by modern climate (informed by all observations) and empirically observed S_R (the local observation). We then explore select case studies of geologic control and suggest process explanations through an analysis of available subsurface geologic and hydrologic field studies.

2 Methods

2.1 Study area

The study area covers the state of California, USA, where three factors make for an ideal setting to explore geologic controls on S_R: i) there is a high diversity of annual precipitation and potential evapotranspiration rates, geologic substrates, and tectonic uplift rates, resulting in large spatial gradients to explore controls on plant biomes and S_R; ii) the local Mediterranean climate (asynchronous seasonal precipitation and energy input, with a long summer dry period) results in almost complete reliance on wet season-replenished storage to sustain evapotranspiration in summer, and iii) existing evidence for widespread and routine use of bedrock water by woody vegetation (McCormick et al., 2021) indicates that water storage capacity inferred from soils databases is insufficient to describe S_R and that bedrock geologic properties that impact plants (nutrients, toxins, and water status) are likely to strongly influence spatial patterns in S_R.

2.2 Identification of lower than climatically expected S_R

To identify locations with a geologic control on S_R, we compare observed S_R to climatically predicted S_R on a per-pixel basis. Locations with an observed S_R lower than expected for the local climate (i.e., low relative to the predicted S_R) are potentially indicative of a geologic limiting factor. The observed S_R is determined based on the previously described approach that records at each location the maximum deficit between cumulative precipitation and cumulative evapotranspiration (Wang-Erlandsson et al., 2016; Dralle et al., 2021), which in California typically exceeds published soils database water storage capacities and must include deeper water storage in bedrock (McCormick et al., 2021). We use a machine learning (random forest) model to predict S_R solely as a function of climatic factors.

2.3 Data sources

All datasets described below previously existed and were ingested and analyzed for this study via the Google Earth Engine cloud computation environment (Gorelick et al., 2017), where spatial joins and spatial resampling were also performed. The data are mapped at the state-wide level in Figure 2.

2.3.1 Observed apparent root-zone water storage capacity, S_R

S_R was calculated following the deficit-based approach described above (see Wang-Erlandsson et al. (2016) for more details), modified to account for the impacts of snow following Dralle et al. (2021). We used the S_R dataset provided by Dralle et al. (2021), which was calculated using data from 2003-2017 and is provided at approximately 1 km pixel resolution. This S_R dataset relies on precipitation data from PRISM (Daly et al., 2015), evapotranspiration data from PML v2 (Y. Zhang et al., 2019), and snow cover from the MODIS Terra normalized difference snow index product (Hall et al., 2010). This S_R dataset also excludes urban areas, open water, and croplands as well as areas in which evapotranspiration exceeded precipitation, which may be due to unaccounted for irrigation, inter-pixel groundwater fluxes or data error.

Consumer dynamics (Kuijper et al., 2015) or episodic disturbances (e.g., fire or logging) may result in lower than climatically possible evapotranspiration and therefore a lower than climatically expected S_R. This is particularly of concern when S_R is inferred from a relatively short timeseries of precipitation and evapotranspiration. Here, because the S_R dataset is inferred from 15 continuous water years, we do not exclude areas with logging or fire. This is motivated by the desire to include as much training data as possible and the finding that paired catchment studies in the region have observed non-detectable changes in dry season streamflow (the time of year when deficits accumulate) just five years after logging (Keppeler & Ziemer, 1990). Spot checking of logged areas indicates that S_R differences between adjacent logged or burned areas during the study period tend to be small relative to differences across geologic contacts or large climate zones.

2.3.2 Climatic predictors of S_R

We used four static climate variables as predictors of S_R:

- Mean annual precipitation, P (mm)
- Mean annual potential evapotranspiration PET (mm)
- The coefficient of variation of annual precipitation, CV_P, equal to the standard deviation of annual precipitation divided by mean annual precipitation.
- The asynchronicity index between precipitation and potential evapotranspiration (in time and in relative magnitude), ASI (Feng et al., 2019)

The precipitation data were obtained from PRISM (Daly et al., 2015) and the potential evapotranspiration data from the MODIS Terra MOD16A2 product (Running et al., 2017) for the period 2003-2017. The ASI raster was previously generated and described in (Hahm, Lapides, et al., 2022). These climate variables were chosen for their widespread availability at relatively high spatial resolution, and because magnitudes and timing of water delivery and water demand are the first order constraints on the amount of water available to plant biomes and the amount that can be taken up by the atmosphere; together P and PET also capture the aridity index (which is important for water partitioning within the classical Budyko framework). The variability of annual precipitation (captured in CV_P) roughly accounts for drought recurrence intervals, which have been hypothesized to be the other primary climatic driver of top-down root zone storage capacity.

2.3.3 Random forest model

We used the RandomForestRegressor module within the scikit-learn Python package (Pedregosa et al., 2011) to predict S_R from four climate variables (mean annual precipitation, the coefficient of variation of annual precipitation, mean annual potential evapotranspiration, and the seasonal asynchronicity between precipitation and energy delivery; detailed in Section 2.3.2). Model accuracy was assessed by first training on a random subset of 75% of the observations and using the resulting preliminary model to predict S_R with the remaining 25% set-aside validation data, after which a final model was trained on the entire dataset. In each case default scikit-learn (version 1.2.0) hyperparameters were used, except for the minimum number of samples per leaf node, which was set to 100 (discussed below).

The choice of random forest modeling over a multiple linear regression approach (with and without interaction terms) is due to the flexibility of the random forest to account for non-linear interactions between climate drivers and S_R, which were were apparent during exploratory data analysis. The choice of random forest modeling over climatic envelope binning approaches is due to the readily available model diagnostics for random forests, specifically feature importance and partial dependence analysis.

The climatically predicted S_R values from a training dataset consisting of an even mixture of climatically optimal and geologically limited S_R will predict the mean of all pixels present in that climate configuration rather than the climatically optimal one. As a result, the extent to which a low S_R for a given climate is indeed low is underestimated, and therefore absolute deviations between observed and predicted S_R should be interpreted as conservative (minimum) estimates.

A concern with any model is overfitting: if all pixels situated within a certain climate configuration identified by the model are geologically rather than climatically limited, the model will not identify them as having lower than climatically expected S_R because no other other pixels with higher S_R for that climate configuration exist. This weakness is unavoidable with both the random forest approach as well as other empirical climatic envelope binning approaches, but can be overcome to some extent by limiting the decision tree depths (i.e., limiting fit) by enforcing a minimum leaf sample size. Specifying tree depth hyperparameters to limit model fitting comes at the potential cost of absolute model accuracy. However, identification of within-climate configuration variability rather than the best predictive accuracy is the overarching goal in this study. Sensitivity explorations indicated that changing tree depth hyperparameters resulted in varying magnitudes of absolute predicted versus observed S_R differences, nevertheless, the spatial patterns highlighted below were robust.

2.3.4 Geologic layers

We compared the output of the random forest model to existing geologic maps. For statewide analyses, we used the 1:750,000 scale digitized Geologic Map of California (Jennings et al., 2010) to interpret patterns in climatically predicted vs. observed S_R. The map was rasterized to 1 km pixel to match the S_R dataset resolution. We additionally used a 1:65,000 scale geologic map (Huber, 1968) to explore km-scale S_R anomaly patterns across granitic plutons in the Sierra Nevada.

3 Results

Our primary findings are that i) while in general climate can predict S_R with reasonable accuracy, there is substantial unexplained variance; ii) regions where observed S_R tends to be lower than climatically predicted are in many cases spatially bounded by geologic contacts, indicative of a bottom-up geologic control on S_R, and iii) these regions of apparent geologic-controlled S_R are not confined to a particular rock type: diverse lithologies—and hypothesized causal mechanisms—are capable of limiting S_R.

3.1 Observed S_R

Over much of the state, S_R falls between 300 and 600 mm (Figure 2e). The largest observed S_R values (yellow areas in Figure 2e) are found along the western flank of the southern Sierra Nevada and the Transverse Ranges, which also have high interannual variability of precipitation (CV_P, Figure 2c) and moderately high energy delivery (PET, Figure 2a). Very low S_R (purple areas in Figure 2e) is observed in the far north-east (Modoc

Plateau), higher elevation regions in the Sierra Nevada, and parts of the foothills surrounding the Sacramento and Central Valley and the Tulare Basin (the large N-S trending region in white in Figure 2 that was masked from analysis primarily due to large agricultural operations and irrigation).

3.2 Climatically predicted S_R

The random forest model driven by the static climate variables predicts S_R with a root mean square error (RMSE) of 132 mm (the average observed S_R of all pixels is 416 mm) and an R^2 of 0.61. This model was specified to have a minimum of 100 leaf nodes to limit the lumping of particular climate configurations within particular geologic units (see above); hyperparameter tuning estimates indicated that the highest accuracy model would have a minimum of 3 leaf nodes but still have an RMSE of 114 mm. In contrast, a multiple linear regression model including interaction terms (not shown) with the same predictor variables achieves an RMSE of 183 mm, much worse than the random forest. At broad scales, the pattern of predicted S_R using the random forest model (Figure 2f) closely resembles the pattern of observed S_R (Figure 2e).

When the final random forest model is trained with all the available data, analysis of feature importance (Figure 3) indicates that CV_P is the most important predictor of S_R, followed by mean annual P. Thus the random forest model indicates that water supply (its inter-annual variability and average magnitude) are the most important climatic controls on S_R within California, with energy supply (PET) and the intra-annual patterns of water and energy delivery (ASI) being less important.

Partial dependence plots (Figure 4) reveal the marginal effect on predictions of S_R to each climate predictor variable. This analysis indicates that high magnitudes of both P and PET and low magnitudes of CV_P predict low values of S_R. S_R increases monotonically with CV_P, whereas the partial dependence of S_R on P exhibits a humped relationship, with a mesic maximum (Good et al., 2017). There is only a weak negative relationship for ASI. We hypothesize that the physical mechanisms behind these patterns are connected to the impacts of annual magnitude and variability of water delivery. S_R is likely low at low P because there is simply not enough precipitation that arrives prior to dry periods to support much evapotranspiration, limiting the size of the deficit (our measure of S_R) that can grow. S_R is similarly low at high P, but for the opposite reason that locations with high P may have their evapotranspiration limited by energy availability (wetter places tend to have lower potential evapotranspiration in California). S_R may increase with CV_P partly because the denominator in that term is P but also because larger relative inter-annual variability means that plants must rely on more stored water to bridge droughts relative to the typical use for the plant community.

3.3 Regions of climatically under-predicted S_R and underlying geology

While the overall patterns of observed and predicted S_R are similar, the differences reveal where geology may limit plant water availability. Figure 2g shows state-wide areas where the observed S_R is less than the climatically predicted S_R. These pixels are, in many regions, strongly clustered in space and include a large N-S trending swath and other smaller regions of the Northern Coast Ranges, the foothills surrounding the north end of the Sacramento Valley, and large parts of the southern Sierra Nevada. While less obvious in the full map of California, the anomalies are spatially organized at local scales as well (Figure 5 a, d, g, j).

The clustering could be due to a regional, systematic top-down disturbance (e.g., fire, logging, or other unaccounted for land-use) or unaccounted-for climate variable in the model. However, comparison of these regions with geologic mapping indicates instead that substrate is playing the primary role in these spatial patterns.

Figure 5 zooms in on four example regions (one for each row) where S_R anomalies roughly coincide in space with mapped geologic units. The left column of Figure 5 shows how pixels with lower than climatically expected S_R (in dark red) tend to be clustered rather than randomly distributed across the landscape, with clusters aligning reasonably well with outlines of geologic formations. The middle column highlights the particular mapped geologic unit whose extent includes areas of anomalous S_R. The right column shows the same mapped geologic unit's outline superimposed on satellite imagery, and Figure 6 more clearly shows these regions to be less forested than their immediate surroundings. The four highlighted regions have distinct rock types (from top to bottom in Figure 5 and clockwise in Figure 6: melange, volcanic, ultramafic, and granitic). The hypothesized mechanisms for geologic control exerted by each of these rock types is explored in the Discussion below.

In Figure 7, we highlight expansive mapped geologic units (more than 1,000 km^2 areal coverage) where the median of the observed minus predicted S_R is less than -20 mm (i.e., geologic units where the observed S_R tends to be substantially less than the climatically predicted S_R across the state of California). We note that i) these substrates span diverse lithologies (including sedimentary, metamorphic and igneous), and that ii) in some cases, the same units identified visually in the regional case-studies (Figure 5) also exhibit anomalously low S_R at the state-wide scale. Overall, 41% of the study area, or approximately 80,000 km^2 had an observed S_R less than -20 mm than the climatically predicted S_R, indicating that roughly a fifth of California's land area may experience geologically limited storage capacity. It is worth noting that Figure 7 identifies young geologic substrates (Quaternary age) as particularly subject to lower than climatically expected storage capacity. This may be due to a variety of mechanisms, including limited time for nutrients to be fixed or mobilized (Chadwick et al., 1999) or water-retaining clay minerals to form (Jefferson et al., 2010).

4 Discussion

To evaluate where geologic substrates may limit biomass or plant productivity and thus water vapor fluxes to the atmosphere, we identify locations where the observed apparent root-zone water storage capacity (S_R) is smaller than expected relative to other locations with similar climate. Similar to empirical ecological approaches that relate plant productivity or biome characteristics to climate, this empirical identification procedure does not determine the mechanisms underlying the lower-than-expected S_R, which could be associated with disturbance, land-use, or herbivory dynamics. The spatial congruence of many of these locations with geologic boundaries, as opposed to e.g. fire or land use boundaries, provides strong evidence for geologic limitations to plant water availability.

4.1 Process-based mechanisms of geologic limitation of S_R

Figure 1 synthesizes previously proposed mechanisms for geologically limited S_R. Two of these mechanisms are hydrologic mechanisms that limit plant-water availability directly (water storage limitation and water excess) whereas the other two mechanisms indirectly limit S_R via chemical processes that limit plant growth (nutrient limitation and toxicity). We stress that these drivers are not necessarily independent: for example, low nutrient availability could limit plants which in turn limits porosity production in the subsurface. Below, we draw on insights from previous field studies to illustrate how these mechanisms operate, using examples revealed by our mapping as illustrative case studies.

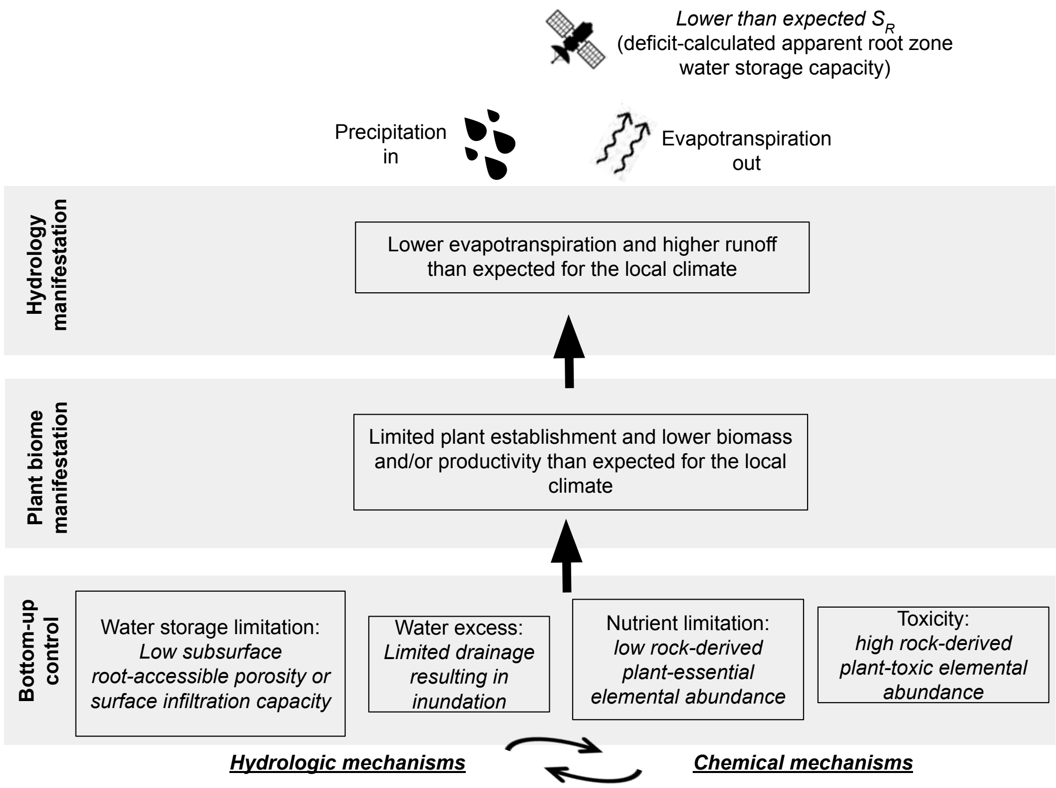

Figure 1. Conceptual diagram illustrating hypothesized geologically mediated controls on apparent root-zone water storage capacity, S_R (lowest row) and corresponding plant biome and hydrologic manifestations. Curved arrows indicate that the geologic controls are not mutually exclusive and may be subject to feedback mechanisms.

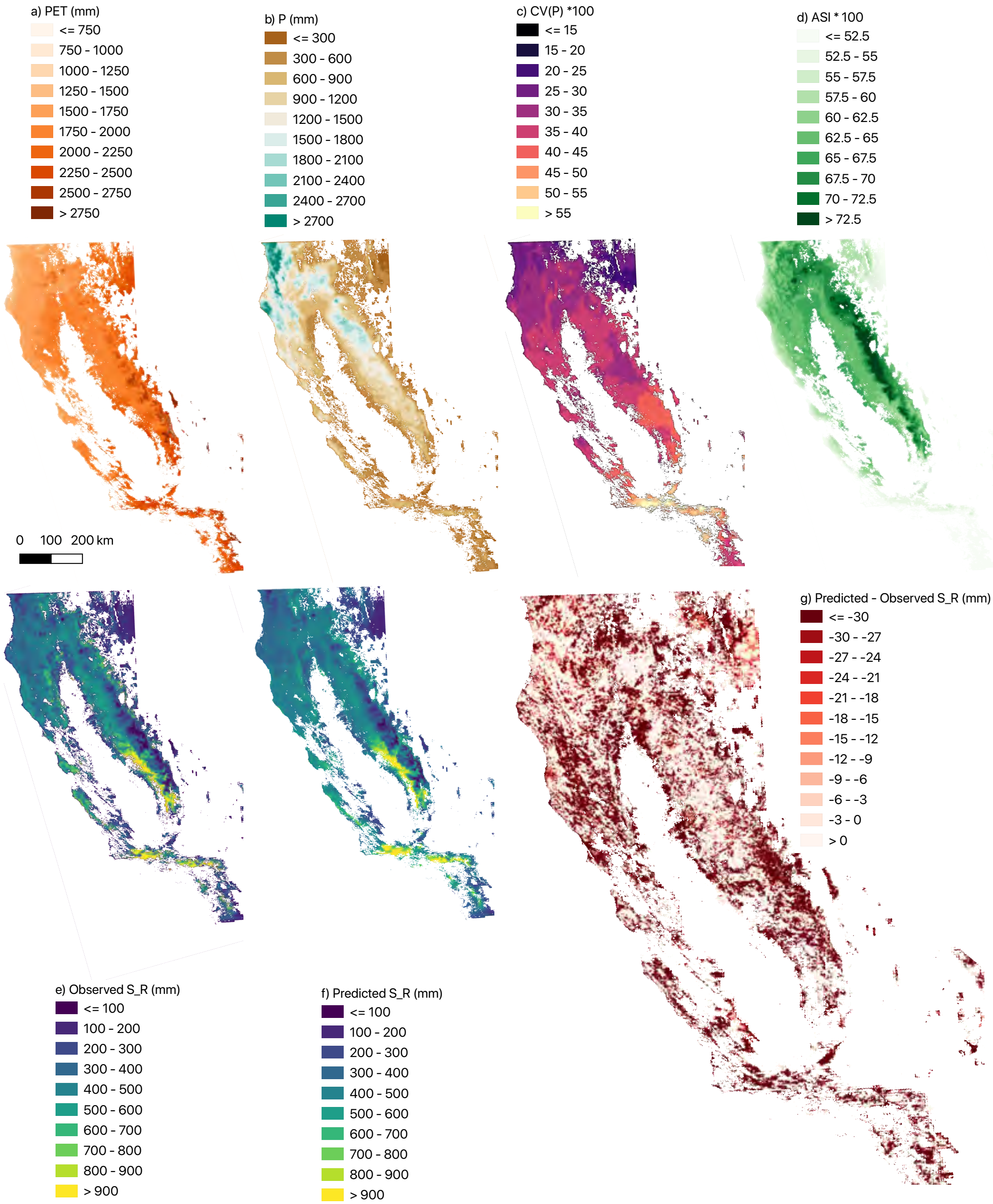

Figure 2. California-wide maps of climatic predictors of S_R (top row) and observed, predicted, and difference between predicted and observed S_R (bottom row). Masked (white) areas are locations where S_R calculation criteria are not met (see Methods).

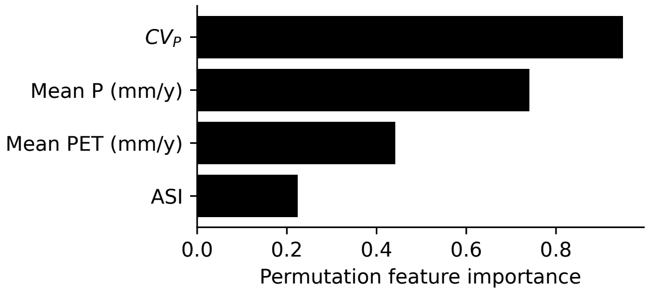

Figure 3. Permutation feature importance of the random forest climate predictors of S_R: higher feature importance indicates that a climate predictor is an important predictor of S_R (inferred by quantifying how much worse the model performs when that variable is randomly shuffled).

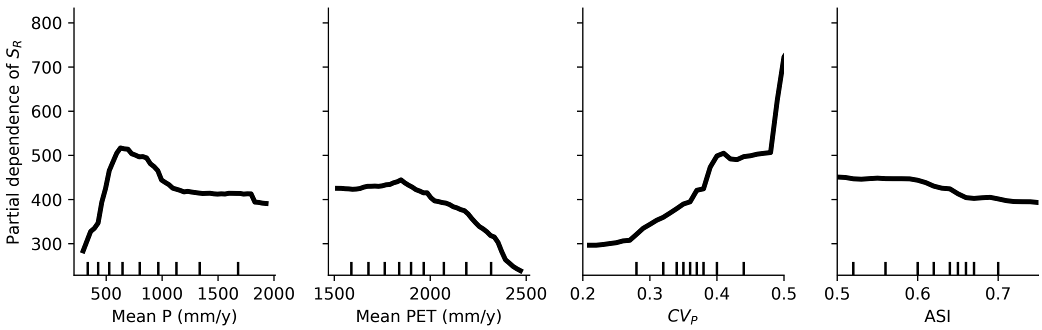

Figure 4. Partial dependence plots show how variation in individual climatic predictor features (x-axes) on average impacts the predicted target variable (S_R, y-axis) when the other climate predictors are controlled for. Vertical lines above x-axes denote decile breaks for the distribution of each climate predictor variable.

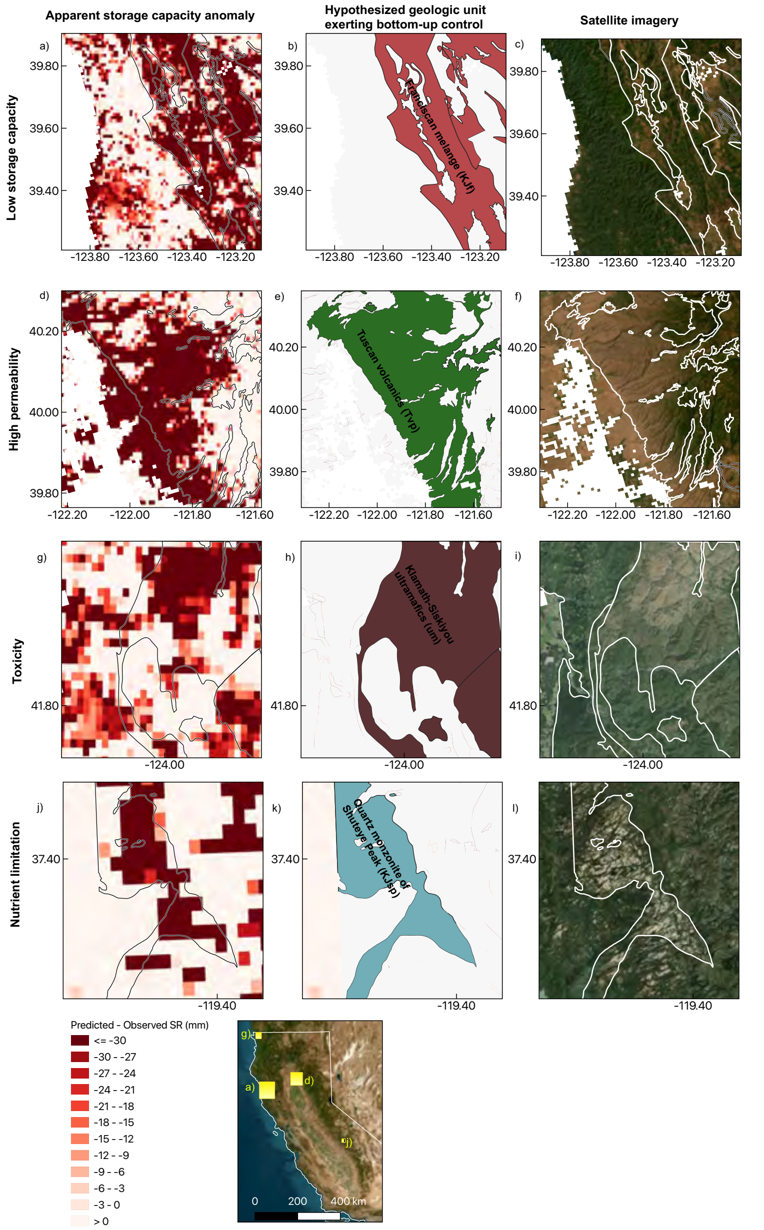

Figure 5. Four regional-scale case studies of apparent geologic control on S_R (one per row). The hypothesized mechanism responsible for anomalously low S_R for the local climate (i.e., red shading in the left column) is identified with the labels at left. The middle column highlights the geologic unit whose spatial extent tends to coincide with a region of anomalous S_R. In the top three rows, the geology mapping comes from the state-wide compilation (Jennings et al., 2010), and in the bottom row from a smaller quadrangle (Huber, 1968). Satellite imagery (from ESRI) in the right column reveals that the low S_R areas also tend to have lower canopy cover than their immediate surroundings. See Discussion for synthesis of prior field studies that support the hypothesized geologic limitation mechanism.

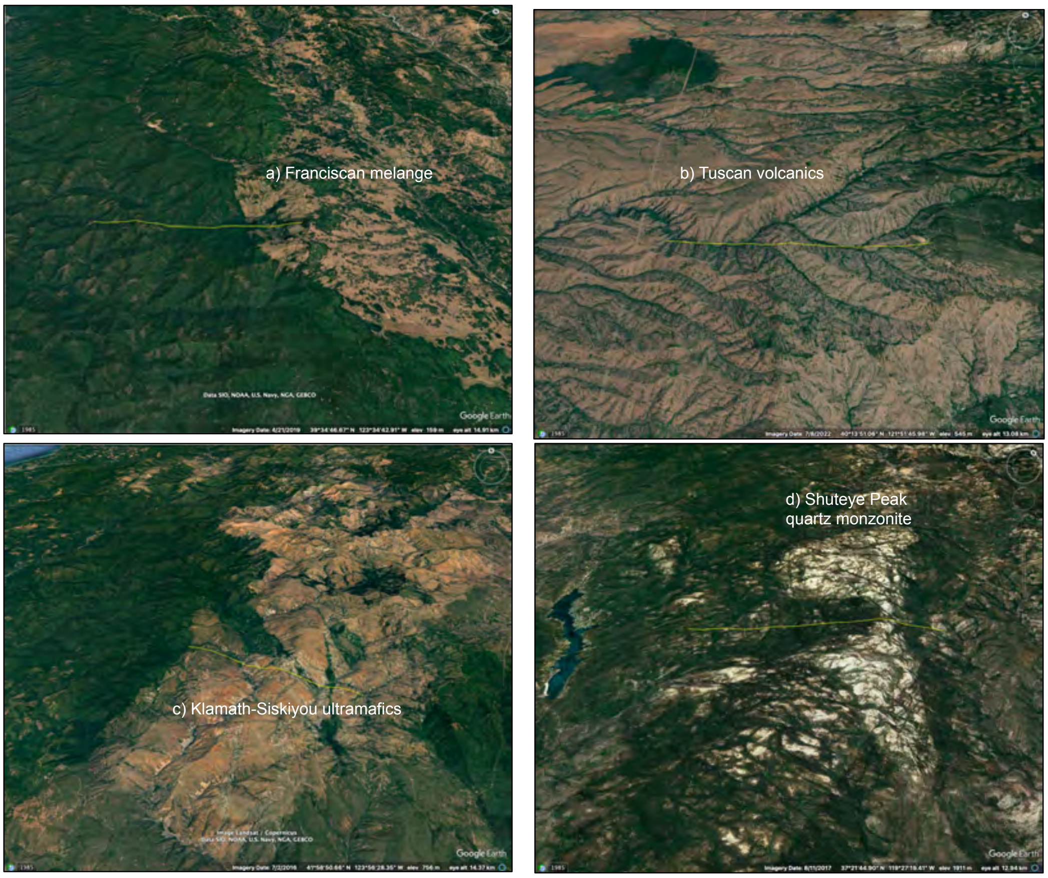

Figure 6. Google Earth imagery with topography of the four case studies highlighted in Figure 5, revealing some of the striking vegetation contrasts over short spatial scales within similar climates that are hypothesized to arise due to geologic controls. The ecotones separating plant communities in these images generally coincide with geologic contacts. In each image, the yellow line is a 10 km scale bar, and the latitude and longitude listed at the lower right of the image is from the center of the scale bar.

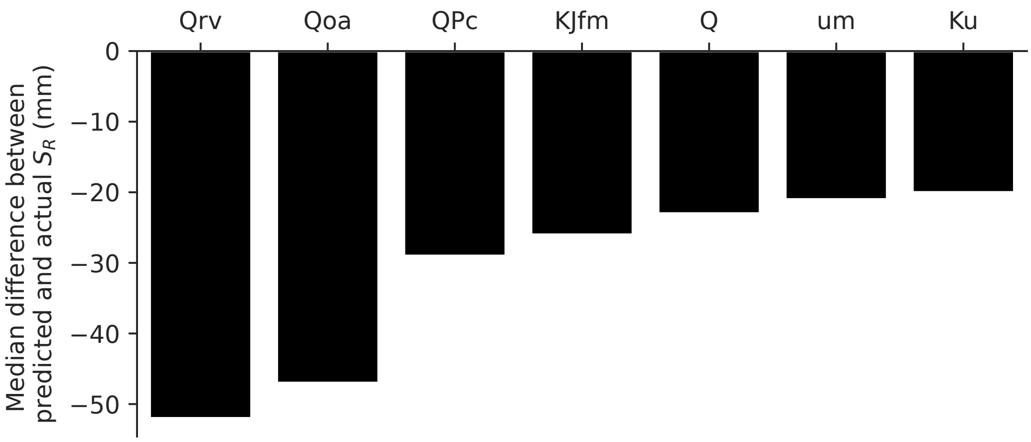

Figure 7. Geologic units with large represented areas ($> 100\ km^2$) that appear to limit root zone storage capacity (i.e., have substantially lower than climatically expected S_R). Key (adapted from Jennings et al. (2010)): Qrv: Volcanic rocks (Holocene) - Recent (Holocene) volcanic flow rocks; minor pyroclastic deposits. Qoa: Marine and nonmarine (continental) sedimentary rocks (Pleistocene) - Older alluvium, lake, playa, and terrace deposits. QPc: Nonmarine (continental) sedimentary rocks (Pleistocene-Holocene) - Pliocene and/or Pleistocene sandstone, shale, and gravel deposits; mostly loosely consolidated. KJfm: Marine sedimentary and metasedimentary rocks (Cretaceous-Jurassic) - Melange of fragmented and sheared Franciscan Complex rocks. Q: Marine and nonmarine (continental) sedimentary rocks (Pleistocene-Holocene) - Alluvium, lake, playa, and terrace deposits; unconsolidated and semi-consolidated. Mostly nonmarine, but includes marine deposits near the coast. um: Plutonic rocks (Mesozoic) - Ultramafic rocks, mostly serpentine. Minor peridotite, gabbro, and diabase; chiefly Mesozoic. Ku: Marine sedimentary and metasedimentary rocks (Upper Cretaceous) - Upper Cretaceous sandstone, shale, and conglomerate.

4.1.1 Water limitation and excess

In both soil and weathered bedrock, connected porosity enables water storage and flow, thereby regulating water status in the root zone (Klos et al., 2018). In upland environments, the weathered bedrock layer is variably thick and typically underlies a physically mobile regolith (soil, in the geomorphological sense) (Rempe & Dietrich, 2014). Weathered bedrock forms as chemical and physical weathering fronts propagate downwards into fresh bedrock as it is nears Earth's surface (Riebe et al., 2017). Under similar climate, spatial gradients in tectonics and lithology can result in variations in weathering extent and thus water storage and flow properties. These variations can result in either limited or excess water, and in some scenarios, both at the same location at different times of year.

For the first case study, we highlight the Central Belt melange of the Franciscan Formation that runs roughly parallel to the coast in the Northern California Coast Ranges (first row in Figure 5). In a region where the local climate can support some of the tallest trees on the planet, the melange is surprisingly sparsely vegetated; instead of the dense forest found immediately to the west, the melange is characterized by an open savanna of deciduous Oregon White oak (*Quercus garryana*) and an herbaceous groundcover (Hahm et al., 2017).

Deep drilling and multiple years of intensive hillslope-scale ecohydrologic field monitoring have resulted in the interpretation that this lower than climatically expected vegetation community arises due the shallow (only 1 to 2 m) propagation of weathering into the fresh melange bedrock (Hahm et al., 2019), which consequently results in limited water storage capacity (about 1/10th of the typical wet season precipitation; in contrast, the Coastal Belt immediately to the west has 20 to 30 m deep weathering fronts and three times greater seasonal water storage (Dralle et al., 2018), with a dense evergreen forest (Figure 6a)). Storage of water from the wet season in the subsurface is critical for plant water supply in the summer dry season in this rain-dominated Mediterranean climate (Hahm, Dralle, et al., 2022). Our mapping in this study extends the insights from hillslope- and catchment-scale field observations and indicates that the melange rock type is associated with lower than climatically expected S_R across the state (the melange is denoted as KJfm in Figure 7).

The low storage capacity of the melange results in both water limitation—in the dry season, when oak pre-dawn water potentials drop below -3 MPa—and water excess, in the wet season, when the subsurface completely saturates repeatedly in storms—resulting in anoxic conditions around flooded roots (Hahm et al., 2018, 2020). The role of excess water as a control on vegetation has also been explored by Sousa et al. (2022); Roebroek et al. (2020); Zipper et al. (2015). The melange presents the interesting situation of rhizosphere water limitation even when a perennially saturated zone is relatively near the surface: in the summer the vadose zone is just a few meters deep, and although the fresh melange beneath is perennially saturated, its extremely low hydraulic conductivity and anoxic conditions apparently prevent root water uptake (Hahm et al., 2020).

In contrast to the scenario where low permeability, perennially saturated fresh bedrock is near the surface, some landscapes can instead have a high conductivity, high porosity substrate that allows infiltrating precipitation to rapidly transit the root zone vertically, draining to deeper aquifers. This form of low vadose zone storage capacity can also lead to water limitation and a lower than climatically expected plant community. These conditions have been documented in karstic terrain in China (H. Liu et al., 2021; Jiang et al., 2020).

We posit that a similar phenomenon may also be possible in highly permeable volcanic bedrock. As a second case study, we highlight a community with low biomass—and low S_R—for the local climate in the Lassen foothills at the north-western end of the

Sacramento Valley (second row in Figure 5). Here, a Pliocene aged volcanic substrate (the Tuscan Formation (Lydon, 1967)) is inhabited by an open oak savanna with abundant rocky outcrops. Both the geomorphology (characterized by buttes) and woody vegetation community (including Interior Live (*Quercus wislizeni* and Blue (*Quercus douglasii*) oaks) are strongly organized along outcrops of particular subhorizontally bedded volcanic deposits (lahars containing tuffs and breccias), as seen in Figure 6b. Based on these bedrock structure and vegetation observations, along with records of high surface infiltration rates and conductivity within permeable beds (Butte County Department of Water and Resource Conservation, 2013), we interpret that in this landscape infiltrating precipitation rapidly transits certain high permeability volcanic beds that comprise the majority of the Formation volumetrically, without significant moisture retention. (Relatively young volcanic landscapes in the Cascades can have relatively little water storage capacity in the near surface and high conductivity (Jefferson et al., 2010; Tague & Grant, 2004) volcanic landscapes). Woody vegetation is minimal on these volcanic beds, but is found along roughly elevation-contour parallel bands where lower conductivity or higher storage capacity beds outcrop at the surface, as vegetation there may experience enhanced water availability from lateral flow or greater retention of infiltrating precipitation.

4.1.2 Toxicity

Toxic concentrations of elements can be released via chemical weathering of underlying bedrock, inhibiting plant growth. Classic examples are associated with ultramafic substrates, and in California there are well-studied examples of high-biodiversity, low-biomass endemic plant communities inhabiting serpentines (Kruckeberg, 1992, 1985; Harrison et al., 2004). In these environments, plants struggle in the presence of exposure to high ratios of Mg:Ca and high Ni (Kruckeberg, 1992).

Consistent with previous observations of low plant biomass on ultramafic substrates, we found that ultramafic areas across the study area tend to have lower than climatically expected S_R (denoted *um* in Figure 7). As a case study, we highlight the dramatic example of a large ultramafic body in the Klamath-Siskiyou region of north-western California—one of the largest in North America (third row in Figure 5 and Figure 6c). This region can climatically support dense evergreen forests, yet the vegetation situated on the serpentine substrate is commonly stunted or altogether absent (Alexander et al., 2007), with scattered individuals of pine, fir and cedar. The inhibited plant growth reduces evapotranspiration, in turn limiting water storage deficits and *apparent* root-zone water storage capacity, as illustrated conceptually in Figure 1. We emphasize that there may in fact be ample water storage capacity, but the stunted plants growing on toxic substrates do not access it, and it is therefore mapped as lower than climatically expected S_R.

4.1.3 Nutrient limitation

Low concentrations of plant-essential nutrients in parent material, low erosion rate and/or high leaching may all contribute to nutrient limitation, stunted vegetation, and lower than expected S_R. In California, nutrient limitation has been associated with ultramafic substrates (see *Toxicity* above), as well as leucogranitic plutons in the Sierra Nevada, where phosphorus concentrations in parent bedrock can be an order of magnitude lower than more mafic adjacent plutons (Hahm et al., 2014). The bottom row of Figure 5 illustrates one such pluton, the Quartz Monzonite of Shuteye Peak, which has low woody plant cover (sparse Jeffrey Pine (*Pinus jeffreyi*)) and large expanses of exposed granitic bedrock, in contrast to nearby granodioritic plutons experiencing a similar climate which are occupied by high biomass evergreen forests, including the charismatic Giant Sequoia (*Sequoiadendron giganteum*); Figure 6d. Ecotones separating the plant communities closely align with mapped intrusive contacts (Huber, 1968; Hahm et al., 2014). Neither Shuteye Peak nor the nearby Bald Mountain were glaciated in the

Pleistocene, and their sparse soil cover has been attributed to nutrient limitation that inhibits root growth which consequently inhibits soil retention (Hahm et al., 2014). This has been hypothesized to result in a feedback cycle that further inhibits weathering and porosity production in the subsurface, which in turn also limits the water storage capacity for trees and their growth (Callahan et al., 2022). Thus, the nutrient and water limitation geologic controls on plant abundance, water use, and ultimately S_R are potentially closely linked via feedback cycles, defining an exciting research frontier.

4.2 Implications for climate change driven plant biome migration and the use of S_R in models

Bioclimatic modeling approaches provide a first approximation to the availability of plant habitat (Pearson & Dawson, 2003). It has long been argued, however, that physiographic, edaphic, and geophysical factors—in addition to climate—should be taken into consideration when predicting and managing for climate change induced species migration (Theobald et al., 2015; Anderson & Ferree, 2010; Hulshof & Spasojevic, 2020; Davis et al., 2018; Butler et al., 2007; Macias-Fauria & Johnson, 2013), a sentiment well captured by Kruckeberg (2013): "given a regional climatic framework, much of the plant species diversity and discontinuity in the region is governed by variations in soil chemistry, and thus by specific variations in the mineralogy of rock substrates." Our work builds on these insights by enabling a direct quantification of the impact of geology over large spatial scales using recently made available, spatially distributed estimates of S_R and a simple, climate-driven machine learning model.

S_R is a key parameter across hydrology, vegetation, and climate models (Seneviratne et al., 2013), because of its large impact on terrestrial water partitioning, plant-water availability and associated carbon uptake, and the associated impacts of latent heat flux and vegetation greenness on the climate. Although previous studies have used both climate and soils databases to establish edaphoclimatic envelopes for modeling vegetation distribution (de Castro Oliveira et al., 2021), there is a growing consensus that traditionally used static soils database derived estimates of S_R are inadequate (Stocker et al., 2023). This is due to the mounting evidence of widespread plant-water uptake from bedrock whose water storage properties are not traditionally included within soils databases (McCormick et al., 2021; Stocker et al., 2023) and because temporally changing vegetation communities can result in shifting magnitudes of S_R at a single location (L. Zhang et al., 2001; Li et al., 2019; Nijzink et al., 2016; Hrachowitz et al., 2021). Our approach offers a path forward for *empirically identifying* geologic limitations on S_R, but we do not see a clear way to *predict* such limitations a priori at large spatial scales at the moment, particularly when they arise due to hydrologic mechanisms (Figure 1). This is due to complicated feedbacks among the various processes and our current inability to directly observe weathering extent and water storage and flow properties at large spatial scales.

4.3 Limitations and future work

The distinction between top-down (climate) versus bottom-up (geologic) drivers of S_R becomes murky over longer time scales. This is partly because landscapes inherit paleoclimate weathering legacies: for example, climate may result in glaciation, which can strip away soil and weathered bedrock, resulting in a proximate bottom-up control on S_R that is facilitated ultimately by a long-term climate history. Climate drivers are also filtered by the subsurface to determine groundwater dynamics, which can strongly impact plant community distribution over individual hillslope lengthscales (Koirala et al., 2017; Roebroek et al., 2020; Fan et al., 2017). Climate also impacts hillslope diffusive and advective erosive processes, which may impact seedling establishment (Toloui-Semnani & Johnson, 2019), the thickness of the weathered bedrock zone and the sizes of colluvial wedges (and potential storage space for water (Ding et al., 2018; Milodowski et al., 2015; Rempe & Dietrich, 2014)) and the spacing of ridges and valleys (Perron et al., 2009). It has also been argued that vegetation 'coevolves' with the subsurface in such a way to produce a particular water storage reservoir: in this view, soils are largely biotic constructs (van Breemen, 1993). The approach outlined in this study is not capable of teasing apart the longer-term connections between top-down and bottom-up drivers of S_R—instead, it takes the current climate at face value and asks whether the empirically observed S_R is lower in some places relative to others with the same climate. While this works in many locations (e.g., the case studies explored above), this empirical approach is incapable of detecting a bottom-up limitation on S_R if all locations for a particular climate are similarly limited by a geologically mediated factor.

An additional complication in identifying bottom-up limitations of S_R can arise in locations with significant inter-pixel lateral groundwater subsidies to vegetation (Roebroek et al., 2020; Fan et al., 2017). In this scenario, a larger than climatically expected S_R may be detected because evapotranspiration is sustained by groundwater flow from elsewhere, which could result in large calculated water storage deficits. We expect this process to be most common at the scale of individual hillslopes, where water that infiltrates near local topographic highs may flow laterally downslope toward local channels. Because the pixel sizes we consider are large relative to local hillslope lengthscales, however, this effect should be minimized in our estimation procedure.

5 Conclusions

We employed a simple machine learning approach to quantify the difference between climatically expected and observed apparent root-zone water storage capacity (S_R). By comparing the resulting patterns with geologic maps, we found strong spatial correspondence between particular substrates and regions of lower than climatically expected S_R. These patterns are indicative of bottom-up controls on the size of the root zone. Our mapping approach is not capable of identifying the mechanisms by which geology limits S_R. However, the patterns we observed are consistent with mechanisms identified in previous field studies, which highlight the role of water availability (excess and limitation), nutrient supply, and toxicity. Although our analysis is not exhaustive, the approach presented here enables extension of hillslope-scale field inferences to much larger areas, and, importantly, does not rely on traditionally used soil water storage capacity databases, which are generally too shallow to capture relevant plant water dynamics in seasonally dry climates. Furthermore, our findings indicate that climate patterns alone can be insufficient predictors of root zone water storage capacity. The subsurface matters, and should be incorporated into earth system models and ecosystem migration management plans in the context of climate change.

6 Open Research

All data sets used in this research were previously published (see references in Methods for details). The Python notebooks used to query Google Earth Engine, aggregate data, and perform the random forest modeling and other data analyses are available on Hydroshare: `https://www.hydroshare.org/resource/be4e3be9e18144908bd4a7baa75a9a4e/` (Hahm, 2023).

Acknowledgments

Funding support from Simon Fraser University, the Canadian Foundation for Innovation, and a Natural Sciences and Engineering Research Council of Canada Discovery Grant.

References

Alexander, E. B., Coleman, R. G., Keeler-Wolfe, T., & Harrison, S. P. (2007). Serpentine plant life of western north america. In *Serpentine geoecology of western north america.* Oxford University Press.

Anderson, M. G., & Ferree, C. E. (2010). Conserving the stage: climate change and the geophysical underpinnings of species diversity. *PloS one*, *5*(7), e11554.

Bouaziz, L. J., Aalbers, E. E., Weerts, A. H., Hegnauer, M., Buiteveld, H., Lammersen, R., … Hrachowitz, M. (2022). Ecosystem adaptation to climate change: the sensitivity of hydrological predictions to time-dynamic model parameters. *Hydrology and Earth System Sciences*, *26*(5), 1295–1318.

Butler, D. R., Malanson, G. P., Walsh, S. J., & Fagre, D. B. (2007). Influences of geomorphology and geology on alpine treeline in the american west—more important than climatic influences? *Physical Geography*, *28*(5), 434–450.

Butte County Department of Water and Resource Conservation. (2013, Feb). *Lower tuscan aquifer monitoring, recharge, and data management project.* Retrieved from `https://www.buttecounty.net/Portals/26/Tuscan/LTAFinalReport.pdf`

Callahan, R. P., Riebe, C. S., Sklar, L. S., Pasquet, S., Ferrier, K. L., Hahm, N. J., W aJnd Taylor, … others (2022). Forest vulnerability to drought controlled by bedrock composition. *Nature Geoscience*, *15*(9), 714–719.

Chadwick, O. A., Derry, L. A., Vitousek, P. M., Huebert, B. J., & Hedin, L. O. (1999). Changing sources of nutrients during four million years of ecosystem development. *Nature*, *397*(6719), 491–497.

Cheng, S., Cheng, L., Qin, S., Zhang, L., Liu, P., Liu, L., … Wang, Q. (2022). Improved understanding of how catchment properties control hydrological partitioning through machine learning. *Water Resources Research*, *58*(4), e2021WR031412.

Daly, C., Smith, J. I., & Olson, K. V. (2015). Mapping atmospheric moisture climatologies across the conterminous united states. *PloS one*, *10*(10), e0141140.

Davis, E. L., Hager, H. A., & Gedalof, Z. (2018). Soil properties as constraints to seedling regeneration beyond alpine treelines in the canadian rocky mountains. *Arctic, Antarctic, and Alpine Research*, *50*(1), e1415625.

Dawson, T. E., Hahm, W., & Crutchfield-Peters, K. (2020). Digging deeper: what the critical zone perspective adds to the study of plant ecophysiology. *New Phytologist*, *226*(3), 666–671.

de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C., & Savenije, H. H. (2016). Influence of soil and climate on root zone storage capacity. *Water Resources Research*, *52*(3), 2009–2024.

de Boer-Euser, T., Palalane, J., Savenije, H., & Juízo, D. (2019). How climate variations are reflected in root zone storage capacities. *Physics and Chemistry of the Earth, Parts A/B/C*, *112*, 83–90.

de Castro Oliveira, G., Arruda, D. M., Fernandes Filho, E. I., Veloso, G. V., Francelino, M. R., & Schaefer, C. E. G. R. (2021). Soil predictors are crucial for modelling vegetation distribution and its responses to climate change. *Science of The Total Environment*, *780*, 146680.

Ding, J., Johnson, E. A., & Martin, Y. E. (2018). Linking soil moisture variation and abundance of plants to geomorphic processes: A generalized model for erosion-uplifting landscapes. *Journal of Geophysical Research: Biogeosciences*, *123*(3), 960–975.

Dralle, D. N., Hahm, W., Chadwick, K. D., McCormick, E., & Rempe, D. M. (2021). Accounting for snow in the estimation of root zone water storage capacity from precipitation and evapotranspiration fluxes. *Hydrology and Earth System Sciences*, *25*(5), 2861–2867.

Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., & Dietrich, W. E. (2018). Quantification of the seasonal hillslope water storage that does not drive streamflow. *Hydrological processes*, *32*(13), 1978–1992.

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C.

(2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, *114*(40), 10572–10577.

Feng, X., Thompson, S. E., Woods, R., & Porporato, A. (2019). Quantifying asynchronicity of precipitation and potential evapotranspiration in mediterranean climates. *Geophysical Research Letters*, *46*(24), 14692–14701.

Gao, H., Hrachowitz, M., Schymanski, S., Fenicia, F., Sriwongsitanon, N., & Savenije, H. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, *41*(22), 7916–7923.

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., & Salvucci, G. (2012). Interdependence of climate, soil, and vegetation as constrained by the budyko curve. *Geophysical Research Letters*, *39*(19).

Good, S. P., Moore, G. W., & Miralles, D. G. (2017). A mesic maximum in biological water use demarcates biome sensitivity to aridity shifts. *Nature ecology & evolution*, *1*(12), 1883–1888.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18–27.

Guswa, A. J. (2008). The influence of climate on root depth: A carbon cost-benefit analysis. *Water Resources Research*, *44*(2).

Guswa, A. J. (2010). Effect of plant uptake strategy on the water- optimal root depth. *Water Resources Research*, *46*(9).

Hahm, W. (2023). *Data and code for: "geologic controls on apparent root-zone storage capacity".* Hydroshare. Retrieved from `https://www.hydroshare.org/resource/be4e3be9e18144908bd4a7baa75a9a4e/`

Hahm, W., Dietrich, W. E., & Dawson, T. E. (2018). Controls on the distribution and resilience of quercus garryana: ecophysiological evidence of oak's water-limitation tolerance. *Ecosphere*, *9*(5), e02218.

Hahm, W., Dralle, D., Lovill, S., Rose, J., Dawson, T., & Dietrich, W. (2017). Exploratory tree survey (2016—eel river critical zone observatory—sagehorn—central belt melange, franciscan complex, northern california coast ranges, usa). hydroshare. *HydroShare https://doi. org/10.4211/hs.7881821a5c0e4ae3822b96a59f4bf8b6*.

Hahm, W., Dralle, D., Sanders, M., Bryk, A., Fauria, K., Huang, M., et al. (2022). Bedrock vadose zone storage dynamics under extreme drought: Consequences for plant water availability, recharge, and runoff. *Water Resources Research*, *58*, e2021WR031781.

Hahm, W., Lapides, D., Rempe, D., McCormick, E., & Dralle, D. (2022). The age of evapotranspiration: Lower-bound constraints from distributed water fluxes across the continental united states. *Water Resources Research*, *58*(10), e2022WR032961.

Hahm, W., Rempe, D., Dralle, D., Dawson, T., & Dietrich, W. (2020). Oak transpiration drawn from the weathered bedrock vadose zone in the summer dry season. *Water Resources Research*, *56*(11), e2020WR027419.

Hahm, W., Rempe, D. M., Dralle, D. N., Dawson, T. E., Lovill, S. M., Bryk, A. B., . . . Dietrich, W. E. (2019). Lithologically controlled subsurface critical zone thickness and water storage capacity determine regional plant community composition. *Water Resources Research*, *55*(4), 3028–3055.

Hahm, W., Riebe, C. S., Lukens, C. E., & Araki, S. (2014). Bedrock composition regulates mountain ecosystems and landscape evolution. *Proceedings of the National Academy of Sciences*, *111*(9), 3338–3343.

Hall, D. K., Riggs, G. A., Foster, J. L., & Kumar, S. V. (2010). Development and evaluation of a cloud-gap-filled modis daily snow-cover product. *Remote sensing of environment*, *114*(3), 496–503.

Harrison, S., Safford, H., & Wakabayashi, J. (2004). Does the age of exposure of ser-

pentine explain variation in endemic plant diversity in california? *International Geology Review*, *46*(3), 235–242.
Holbrook, W. S., Riebe, C. S., Elwaseif, M., L. Hayes, J., Basler-Reeder, K., L. Harry, D., ... W. Hopmans, J. (2014). Geophysical constraints on deep weathering and water storage potential in the southern sierra critical zone observatory. *Earth Surface Processes and Landforms*, *39*(3), 366–380.
Hrachowitz, M., Stockinger, M., Coenders-Gerrits, M., van der Ent, R., Bogena, H., Lücke, A., & Stumpp, C. (2021). Reduction of vegetation-accessible water storage capacity after deforestation affects catchment travel time distributions and increases young water fractions in a headwater catchment. *Hydrology and earth system sciences*, *25*(9), 4887–4915.
Huber, N. K. (1968). *Geologic map of the shuteye peak quadrangle, sierra nevada, california.* Retrieved from https://catalog.data.gov/dataset/digital-geologic-gis-map-of-the-shuteye-peak-15-quadrangle-california-nps-grd-gri-yose-shu
Hulshof, C. M., & Spasojevic, M. J. (2020). The edaphic control of plant diversity. *Global Ecology and Biogeography*, *29*(10), 1634–1650.
Jefferson, A., Grant, G., Lewis, S., & Lancaster, S. (2010). Coevolution of hydrology and topography on a basalt landscape in the oregon cascade range, usa. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *35*(7), 803–816.
Jennings, C. W., Gutierrez, C., Bryant, W., Saucedo, G., & Wills, C. (2010, Dec). *Geologic map of california version 2.0.* Retrieved from https://maps.conservation.ca.gov/cgs/#datalist
Jiang, Z., Liu, H., Wang, H., Peng, J., Meersmans, J., Green, S. M., ... Song, Z. (2020). Bedrock geochemistry influences vegetation growth by regulating the regolith water holding capacity. *Nature communications*, *11*(1), 2392.
Keppeler, E. T., & Ziemer, R. R. (1990). Logging effects on streamflow: water yield and summer low flows at caspar creek in northwestern california. *Water resources research*, *26*(7), 1669–1679.
Kleidon, A., & Heimann, M. (1998). A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle. *Global Change Biology*, *4*(3), 275–286.
Klos, P. Z., Goulden, M. L., Riebe, C. S., Tague, C. L., O'Geen, A. T., Flinchum, B. A., ... others (2018). Subsurface plant-accessible water in mountain ecosystems with a mediterranean climate. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1277.
Koirala, S., Jung, M., Reichstein, M., de Graaf, I. E., Camps-Valls, G., Ichii, K., ... others (2017). Global distribution of groundwater-vegetation spatial covariation. *Geophysical Research Letters*, *44*(9), 4134–4142.
Kruckeberg, A. (1985). *California serpentines: flora, vegetation, geology, soils, and management problems.* Univ of California Press.
Kruckeberg, A. (1992). Plant life of western north american ultramafics. *The ecology of areas with serpentinized rocks: a world view*, 31–73.
Kruckeberg, A. (2013). Soil diversity and the distribution of plants, with examples from western north america. *Madroño*, *60*(4), 267–292.
Kuijper, D. P., Te Beest, M., Churski, M., & Cromsigt, J. (2015). Bottom-up and topdown forces shaping wooded ecosystems: lessons from a cross-biome comparison. *Trophic Ecology. Cambridge University Press, Cambridge*, 107–133.
Lapides, D. A., Hahm, W., Rempe, D. M., Whiting, J., & Dralle, D. N. (2022). Causes of missing snowmelt following drought. *Geophysical Research Letters*, *49*(19), e2022GL100505.
Li, H., Si, B., Wu, P., & McDonnell, J. J. (2019). Water mining from the deep critical zone by apple trees growing on loess. *Hydrological Processes*, *33*(2), 320–327.

Liu, H., Dai, J., Xu, C., Peng, J., Wu, X., & Wang, H. (2021). Bedrock-associated belowground and aboveground interactions and their implications for vegetation restoration in the karst critical zone of subtropical southwest china. *Progress in Physical Geography: Earth and Environment*, *45*(1), 7–19.

Liu, M., Zhang, B., & He, X. (2022). Climate rather than vegetation changes dominate changes in effective vegetation available water capacity. *Water Resources Research*, *58*(3), e2021WR030319.

Lydon, P. A. (1967). The origin of the tuscan buttes and the volume of the tuscan formation in northern california. *Short Contributions to California Geology: Special Report*, *91*, 17–26.

Macias-Fauria, M., & Johnson, E. A. (2013). Warming-induced upslope advance of subalpine forest is severely limited by geomorphic processes. *Proceedings of the National Academy of Sciences*, *110*(20), 8117–8122.

McCormick, E. L., Dralle, D. N., Hahm, W., Tune, A. K., Schmidt, L. M., Chadwick, K. D., & Rempe, D. M. (2021). Widespread woody plant use of water stored in bedrock. *Nature*, *597*(7875), 225–229.

McLaughlin, B. C., Blakey, R., Weitz, A. P., Feng, X., Brown, B. J., Ackerly, D. D., ... Thompson, S. E. (2020). Weather underground: Subsurface hydrologic processes mediate tree vulnerability to extreme climatic drought. *Global change biology*, *26*(5), 3091–3107.

Milly, P. (1994). Climate, interseasonal storage of soil water, and the annual water balance. *Advances in Water Resources*, *17*(1-2), 19–24.

Milodowski, D. T., Mudd, S. M., & Mitchard, E. T. (2015). Erosion rates as a potential bottom-up control of forest structural characteristics in the sierra nevada mountains. *Ecology*, *96*(1), 31–38.

Morford, S. L., Houlton, B. Z., & Dahlgren, R. A. (2011). Increased forest ecosystem carbon and nitrogen storage from nitrogen rich bedrock. *Nature*, *477*(7362), 78–81.

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., ... others (2016). The evolution of root-zone moisture capacities after deforestation: a step towards hydrological predictions under change? *Hydrology and Earth System Sciences*, *20*(12), 4775–4799.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global ecology and biogeography*, *12*(5), 361–371.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perron, J. T., Kirchner, J. W., & Dietrich, W. E. (2009). Formation of evenly spaced ridges and valleys. *Nature*, *460*(7254), 502–505.

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *The American Naturalist*, *164*(5), 625–632.

Rempe, D. M., & Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy of Sciences*, *111*(18), 6576–6581.

Riebe, C. S., Hahm, W., & Brantley, S. L. (2017). Controls on deep critical zone architecture: A historical review and four testable hypotheses. *Earth Surface Processes and Landforms*, *42*(1), 128–156.

Roebroek, C. T., Melsen, L. A., Hoek van Dijke, A. J., Fan, Y., & Teuling, A. J. (2020). Global distribution of hydrologic controls on forest growth. *Hydrology and Earth System Sciences*, *24*(9), 4625–4639.

Running, S., Mu, Q., & Zhao, M. (2017). Mod16a2 modis/terra net evapotranspiration 8-day l4 global 500m sin grid v006. *NASA EOSDIS Land Processes DAAC*, *6*.

Schenk, H. J. (2008). The shallowest possible water extraction profile: A null model for global root distributionsall rights reserved. no part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. *Vadose Zone Journal*, *7*(3), 1119–1124.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., & Hutley, L. B. (2008). An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and earth system sciences*, *12*(3), 913–932.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., ... others (2013). Impact of soil moisture-climate feedbacks on cmip5 projections: First results from the glace-cmip5 experiment. *Geophysical Research Letters*, *40*(19), 5212–5217.

Singh, C., Wang-Erlandsson, L., Fetzer, I., Rockström, J., & Van Der Ent, R. (2020). Rootzone storage capacity reveals drought coping strategies along rainforest-savanna transitions. *Environmental Research Letters*, *15*(12), 124021.

Sousa, T. R., Schietti, J., Ribeiro, I. O., Emílio, T., Fernández, R. H., ter Steege, H., ... others (2022). Water table depth modulates productivity and biomass across amazonian forests. *Global Ecology and Biogeography*.

Speich, M. J., Lischke, H., & Zappa, M. (2018). Testing an optimality-based model of rooting zone water storage capacity in temperate forests. *Hydrology and Earth System Sciences*, *22*(7), 4097–4124.

Stocker, B. D., Tumber-Dávila, S. J., Konings, A. G., Anderson, M. C., Hain, C., & Jackson, R. B. (2023). Global patterns of water storage in the rooting zones of vegetation. *Nature Geoscience*, 1–7.

Tague, C., & Grant, G. E. (2004). A geological framework for interpreting the low-flow regimes of cascade streams, willamette river basin, oregon. *Water Resources Research*, *40*(4).

Theobald, D. M., Harrison-Atlas, D., Monahan, W. B., & Albano, C. M. (2015). Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PloS one*, *10*(12), e0143619.

Toloui-Semnani, M., & Johnson, E. A. (2019). Processes that cause natural unforested areas in canadian rockies below temperature treeline. *Geomorphology*, *347*, 106857.

Trautmann, T., Koirala, S., Carvalhais, N., Güntner, A., & Jung, M. (2022). The importance of vegetation in understanding terrestrial water storage variations. *Hydrology and Earth System Sciences*, *26*(4), 1089–1109.

van Breemen, N. (1993). Soils as biotic constructs favouring net primary productivity. *Geoderma*, *57*(3), 183–211.

van Oorschot, F., van der Ent, R. J., Hrachowitz, M., & Alessandri, A. (2021). Climate-controlled root zone parameters show potential to improve water flux simulations by land surface models. *Earth System Dynamics*, *12*(2), 725–743.

Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I., ... Savenije, H. H. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrology and Earth System Sciences*, *20*(4), 1459–1481.

Witty, J. H., Graham, R. C., Hubbert, K. R., Doolittle, J. A., & Wald, J. A. (2003). Contributions of water supply from the weathered bedrock zone to forest soil quality. *Geoderma*, *114*(3-4), 389–400.

Yang, Y., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, *52*(10), 8260–8276.

Zhang, C., Wang, Y., Jia, X., An, Z., et al. (2020). Variations in capacity and storage of plant-available water in deep profiles along a revegetation and precipitation gradient. *Journal of Hydrology*, *581*, 124401.

Zhang, L., Dawes, W., & Walker, G. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water resources research*, *37*(3), 701–708.

Zhang, Y., Kong, D., Gan, R., Chiew, F. H., McVicar, T. R., Zhang, Q., & Yang, Y. (2019). Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote sensing of environment*, *222*, 165–182.

Zhu, X., Liu, H., He, W., Wu, L., & Liu, F. (2023). Regolith water storage patterns determine vegetation productivity in global karst regions. *Geoderma*, *430*, 116292.

Zipper, S. C., Soylu, M. E., Booth, E. G., & Loheide, S. P. (2015). Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. *Water Resources Research*, *51*(8), 6338–6358.