Root Foraging Alters Global Patterns of Ecosystem Legacy From Climate Perturbations

M. Berkelhammer, B. Drewniak, B. Ahlswede, and M. A. Gonzalez-Meler

1Department of Earth and Environmental Sciences, University of Illinois at Chicago, Chicago, IL, USA, 2Argonne National Lab, Lemont, IL, USA

Abstract The response of terrestrial ecosystems to climate perturbations typically persist longer than the timescale of the forcing, a phenomenon broadly referred to as legacy. Understanding the strength of legacy is critical for predicting ecosystem sensitivity to climate extremes and the extent that persistent changes in surface-atmosphere exchange might feedback onto climate. The cause of ecosystem legacy has been associated with myriad factors such as changes in aboveground biomass, however, few studies have tested how changes in the depth distribution of fine roots in response to perturbation might alter an ecosystem’s recovery time. We explore this question using an Earth System Model that includes a dynamic root module where vegetation can forage for water and nutrients by altering their root profiles. The global simulations presented here show that in response to water stress events most ecosystems deepen their root profiles. In semi-arid ecosystems, this response expedites recovery (i.e., less legacy) relative to simulations without dynamics roots because access to deeper water pools after the initial event remains favorable. In wetter ecosystems, the development of deeper root profiles slows down the recovery timescale (i.e., more legacy) because the deeper root profile reduces access to nutrients and is thus unfavorable. The simulations show that while the inclusion of dynamic roots might only minimally affect global patterns of Gross Primary Productivity and Evapotranspiration, the shift in root profile alters the timescale of recovery. Studies interested in the sustained response of land surfaces fluxes to climate disturbances should consider models that include dynamic root capability.

Plain Language Summary When vegetation is stressed from climate events such as drought, they alter the depths where roots are developed. This allows plants to better access resources that have become limited—typically water. Most Earth System Models do not include this dynamic and instead use a fixed root profile. It is unclear how the absence of dynamic roots affects the skill at which these models capture the response and recovery to climate stress events. Here, we use a climate model with dynamic roots to see how the presence of foraging roots affects the way different ecosystems respond to drought. We show that in drier ecosystems, vegetation develop deeper roots after drought and this allows the vegetation to recover more quickly from drought. Wetter ecosystems also develop deeper roots in response to stress, however, this actually slows down their rate of recovery. Ecosystem response to drought is a critical component of climate models and we show that models that lack dynamic roots miss a key component of terrestrial ecosystem recovery from stressful climate events.

1. Introduction

Ecosystems across all global biomes display some varying level of sensitivity to antecedent conditions. Key ecosystem processes such as transpiration, CO₂ flux and shortwave absorption are thus not only instantaneously responding to external conditions but display a response that integrates the conditions over recent hours, days or years (Ogle et al., 2015). This is a coupled climate-ecosystem phenomenon referred to as “legacy” or “memory” that emerges from internal ecosystem dynamics and is distinct from the persistence of climate or weather patterns that might arise from lower frequency ocean-atmosphere modes such as the Pacific Decadal Oscillation or El Niño Southern Oscillation (Bunde et al., 2013; Kumar et al., 2019). While legacy effects have been observed and classified by ecologists for decades, it is a topic that has received renewed interest recently owing both to concerns that the effects of more frequent extreme events could be compounded by persistent legacy leading to mortality or bifurcation (L. D. Anderegg, et al., 2013; McDowell et al., 2013; Szejner et al., 2020; Trugman et al., 2018) and the recognition that the predictive skill of models is limited without considering these effects (W. R. Anderegg, et al., 2015; Kolus et al., 2019). With respect to the latter, the lack of realistic legacy effects in land surface models is problematic because Earth System Models (ESMs) exclude feedbacks that might act
From the standpoint of climate feedbacks, ecosystem legacy effects can involve a variety of processes such as direct changes in carbon (Gross Primary Production-GPP- or respiration), water (transpiration), energy fluxes or shifts in aboveground biomass that affect frictional and radiative properties of the land surface (Dewar et al., 1994; Galiano et al., 2011; Liu et al., 2019; van der Molen et al., 2011). The various observational tools that have been utilized to quantify ecosystem legacy such as tree ring records (Gazol et al., 2020; Kannenberg et al., 2019; Peltier & Ogle, 2020), eddy covariance (Liu et al., 2019) and satellite retrievals (Kolus et al., 2019; Seddon et al., 2016) have distinct sensitivities to these different legacy components (Ogle et al., 2015). For example, tree ring records are a direct indicator of carbon allocated to woody biomass, eddy covariance provides information on land surface-atmosphere exchange whereas satellite retrievals of emissivity (e.g., MODIS) or canopy structure (e.g., LiDAR) are sensitive to canopy physical and radiative properties. While these approaches provide overlapping information (e.g., tree ring growth is partially controlled by carbon uptake [Campioli et al., 2016]), the different approaches also yield seemingly disparate information. For example, Kannenberg et al. (2019) and Gazol et al. (2020) both note that sustained reductions in tree ring widths following drought were not mirrored by persistent reductions in satellite indices of greenness or plot-scale GPP. These differences could reflect distinct components of legacy or biases in the observational tools. For example, satellite retrievals may lack sufficient sensitivity to capture subtle legacy effects or simply miss changes in below-canopy dynamics. Reconciling this information to develop a clearer picture of the timescales and strength of legacy within different components of ecosystems is key for diagnosing missing legacy dynamics in land surface models.

By taking advantage of the different observational and statistical approaches as well as efforts to experimentally perturb ecosystems (Belk et al., 2007; Gonzalez-Valencia et al., 2014; Herzog et al., 2014; Jiang et al., 2019; Liu et al., 2019; Ogle et al., 2015; Zweifel et al., 2020), a clearer picture has begun to emerge about mechanisms driving legacy (Kannenberg et al., 2020; Monger et al., 2015; Ovenden et al., 2021). These mechanisms include changes in allocation between and within above- (leaf, wood, and stem) and belowground (roots) carbon pools (Phillips et al., 2016; Zweifel et al., 2020), damage to organs such as embolisms in xylem tissue or reduced stomatal control (McDowell et al., 2013), exhaustion or buildup of stored carbon pools (Richardson et al., 2015), changes in the water table depth or deep soil moisture (Amenu et al., 2005; Kumar et al., 2019; O. Sala et al., 1992) and shifts in vulnerability to pests or pathogens that can extend the effects of an isolated climate event (Flower & Gonzalez-Meler, 2015). Shifting allocation patterns between leaves, woody biomass and fine roots has emerged as a critical and ubiquitous source of legacy. This effect can be predicted from ecological theory on optimizing resources (water, nutrients and light) to maximize profits (carbon pools) (Bloom et al., 1985; McCarthy & Enquist, 2007; McNickelle et al., 2016; Poorter et al., 2012; Thornley, 1998). It follows that drought stress—the most widespread global cause of seasonal to interannual ecosystem stress (Seddon et al., 2016)—leads to both reduced total carbon pools and increased investment of those pools belowground to forage for water (Brunner et al., 2009; Drewniak, 2019; Joslin et al., 2000; Lu et al., 2019; Markewitz et al., 2010; Metcalfe et al., 2008). The loss of carbon dedicated to leafy material in exchange for root mass leads to a persistent multi-year legacy on primary productivity and consequently resources available to rebuild the canopy (Galiano et al., 2011; Zweifel et al., 2020). On the other hand, drought stress may not change carbon allocated to leaves and instead shift allocation only between wood and fine roots and thus have a minor impact on aboveground greenness but have an extended impact on belowground processes and, consequently, access to nutrients and water (Doughty et al., 2014; Dybzinski et al., 2011; Phillips et al., 2016).

We focus hereafter on this question of how transient changes in the depth distribution of fine roots following stress events influences temporal patterns in ecosystem productivity. Although some observational studies have supported theoretical predictions where stress promotes shifting carbon investment belowground (Canadell et al., 1996; Markesteijn & Poorter, 2009; Schenk & Jackson, 2002), this effect is not consistently borne out and does not necessarily predict how a change in investment in root systems influences ecosystem recovery or legacy. On the one hand, the investment in deeper roots following drought could expedite recovery if the modified root distribution proves beneficial. For example, an initial dry period may lead to a persistent reduction in root zone
moisture or a drop in the water table (Kumar et al., 2019; Monger et al., 2015; van der Molen et al., 2011). In this case, the investment in a deeper fine root pool would expedite recovery and reduce legacy by “anticipating” sustained water stress. On the other hand, the initial response to forage for water with deeper roots may slow down the recovery if the root profile is now “maladjusted” (Zweifel et al., 2020) and the reduced access to shallow water and nutrients delays recovery. In this case, legacy of a drought could be extended by the change in root structure. Over time, periodic droughts may encourage the development of root systems that are optimized to reduce vulnerability to extreme water stress events. This can generate a circumstance where during periods when water is not limited, the root systems are suboptimal due to the response to prior stress events (i.e., life history). Under the circumstance where water is plentiful but the root profile is optimized for drought, the root system may shallow enabling greater access to nutrients and near surface soil moisture that is available at less hydraulic cost. The shallower root system either allows the ecosystem to remain highly productive and extend the legacy of the bountiful times or leave the system vulnerable to future periods of water limitation especially if moisture limitation is the typical state of the system (Jiang et al., 2019). Having a plastic root structure could thus have the effect of either increasing or decreasing the strength of legacy but how these responses manifest across global bioclimatic gradients has not yet been tested (Phillips et al., 2016).

The response of root systems to an exogenous forcing like drought is itself complicated to predict (Hendrick & Pregitzer, 1996; Metcalfe et al., 2008) and whether this change enhances or diminishes legacy is largely unknown (Brunner et al., 2009; Phillips et al., 2016). Observational studies show that the response of roots to climate forcing and the cascading effects these changes have on productivity are a function of soil structure and nutrient availability, plant type, severity and length of the climate anomaly and background climate state making it difficult to scale up from local studies (Doughty et al., 2014; Germon et al., 2020; Kou et al., 2018; Matamala et al., 2003; McCormack et al., 2015; Warren et al., 2015). One way to develop more universal hypotheses on the role of roots in generating ecosystem legacy is to explore how legacy in land surface models is affected by changing root profiles. Land Surface Models that include dynamic roots have been developed (Bouda & Saiers, 2017; Drewniak, 2019; El Masri et al., 2015; Lu et al., 2019; Sakschewski et al., 2021) and these can be implemented into ESMs and forced with a broad spectrum of climate inputs across global biomes. For example, Drewniak (2019) recently implemented a dynamic root module into the Energy Exascale Earth System Model (E3SM) and reproduced some of the key features observed in root profile measurements including a consistent deepening of roots in response to water stress. While the model could not reproduce root profiles in some ecosystems, such as the dimorphic pattern of roots in the dry tropics (Sakschewski et al., 2021), the addition of dynamic roots did modestly improve global estimates of productivity. Similarly, Lu et al. (2019) developed a 3-dimensional dynamic root module that allowed root systems to develop from a null state. They also reproduced similar dynamic deepening of roots in response to reduced water availability and simulated the evolution of root profiles as forest stands age consistent with data from chronosequences. While these models are limited in their ability to simulate root morphology, root phenology, dynamic allocation, response to stress and soil moisture-groundwater interactions they, nonetheless, provide a framework to consider questions of how foraging patterns in roots alter ecosystem legacy strength and timescale across biomes.

In this paper, we utilize a pair of global ESM simulations forced with historical climate in both a default mode (fixed root profiles i.e., control) and with the inclusion of a dynamic root profile module (Drewniak, 2019) to explore the impact of root dynamics on ecosystem response to perturbation. In Section 1 of the study, we assess how well the control simulations reproduce the legacy of GPP with respect to satellite observations. In Section 2, we classify how root systems respond to stress using a clustering algorithm to identify dominant global patterns in root response to perturbation across biomes and plant functional types. In Section 3, we assess the extent to which the different responses of root profile to perturbation either enhanced legacy or whether the altered root profile led to a more optimal state that expedited recovery. While the goals of this paper are first to develop new hypotheses on the impact of dynamic roots on temporal patterns in ecosystem fluxes, we also explore the fundamental question of when—or if—the inclusion of dynamics roots should be a default process in land surface models. It might seem intuitive that a documented process such as root foraging should be necessary for a modeled ecosystem to meet its water demands, especially during periods of water stress. However, as soil layers with high densities of roots dry out, water from moister soil layers will flow along the matric potential gradient through roots towards these layers (Jarvis, 2011). In models with fixed root profiles, water flows towards the roots instead of the roots migrating towards the water. This is a physically-based emergent feature of land surface models that enables vegetation to meet water demands without requiring the additional complexity of dynamic roots. This is to say,
that adding dynamics roots—or any new ecological process to an ESM—should not necessarily be done simply because it can be done especially if the missing process is being compensated for through other dynamics. New processes add computational costs and can add uncertainty that might reduce the utility of the model especially if the relevant ecological process is not that well understood (Kyker-Snowman et al., 2021). We therefore conclude the paper by posing the question of whether dynamic roots should be a default configuration in the context of recent criteria presented by Moore (2022).

2. Methods

2.1. Model Simulations

All E3SM simulations were performed using the E3SM Land Model (ELM) v1 version (tag 78d0813) (Golaz et al., 2019) in an offline mode using atmospheric forcing from the Global Soil Wetness Project 3 (GSWP3; (Dirmeyer et al., 2002)) at a resolution of 0.5° by 0.5° and a timestep of 1 hr. The biogeochemistry mode of ELM is set to the default Converging Trophic Cascade (see: Burrows et al. (2020)), which includes carbon-nitrogen coupling and nitrification/denitrification processes. However, phosphorus limitation was not included in the simulations. Prognostic crops were turned on, but irrigation was not used. Model spinup was performed by cycling the GSWP3 over the years 1901–1920 for 300 years using the accelerated spinup mode and again for an additional 300 years in regular spinup mode until reaching steady state according to the spinup procedures recommended by Koven et al. (2013) and Thornton and Rosenbloom (2005). Following spinup, all simulations were run over the historical period from 1901 to 2010 with the GSWP3 atmospheric forcing data and a constant CO₂ value of 367 ppm. We performed multiple simulations with various configurations of water stress and root turnover rates but due to the similarity of the simulations, we report outputs from two simulations referred to hereafter as No Dyn. or control and Dyn. The No Dyn. simulation is the control run with the default ELM configuration without dynamic roots. In this configuration, plant rooting depth and distribution are fixed following profiles described in Zeng (2001). The Dyn. simulation was performed with dynamic roots turned on and with an increase in the weight given to water in the root foraging scheme by setting $f_{\text{to}} \leq 0.5$ as in Drewniak (2019). This does not affect the actual water stress in the vegetation but reduces the minimum preference for roots to seek soil layers with water to facilitate more foraging dynamics. Within the Dyn. simulation we also increased the root turnover time from 1 to 2 years to enhance root legacy but this had virtually no impact on the results compared with simulations run with turnover time kept at default (see Supplementary Material in Supporting Information S1).

From the historical E3SM simulations, we extract monthly-averaged GPP, transpiration, relative fine root fraction per soil layer, temperature and precipitation as well as the distribution of Plant Functional Types (PFTs) from each grid cell (Table 1). As noted above, studies on legacy have relied on a wide variety of metrics ranging from tree ring growth, canopy structure or net ecosystem exchange of carbon (Kannenberg et al., 2020; Liu et al., 2019; O. E. Sala et al., 2012). We utilize model outputs of GPP and its persistence following perturbation as the ecosystem metric to track legacy because GPP is closely linked with the overall ecosystem carbon pool available to recover from (or extend) an event and also that satellite estimates of GPP from solar-induced fluorescence (SIF) and near infrared reflectance (Section 2.2) provide a global-scale and independent benchmark to compare simulated legacy timescales against. After extracting the GPP data from the simulations, we apply a PFT filter to remove grid cells that do not have a dominant PFT—defined here as grid cells where a single PFT accounts for ≥50% of the grid area. The reason for this filtering was to balance the need to use grid-averaged data to facilitate comparisons with the satellite data while also ensuring each grid cell can be associated with a specific PFT to facilitate analysis of PFT-specific dynamics. For each grid cell, we generated monthly GPP anomalies by detrending each

| PFT                                      | Numeric | Acronym |
|------------------------------------------|---------|---------|
| Needleaf Evergreen Temperature Tree      | 2       | NETT    |
| Needleaf Evergreen Boreal Tree           | 3       | NEBT    |
| Needleaf Deciduous Boreal Tree           | 4       | NDBT    |
| Broadleaf Evergreen Tropical Tree        | 5       | BETT    |
| Broadleaf Evergreen Temperate Tree       | 6       | BETT    |
| Broadleaf Deciduous Tropical Tree        | 7       | BDTrT   |
| Broadleaf deciduous temperate tree       | 8       | BDTT    |
| Broadleaf Deciduous Boreal Tree          | 9       | BDDBT   |
| Broadleaf Evergreen Shrub                | 10      | BES     |
| Broadleaf Deciduous Temperate Shrub      | 11      | BDTS    |
| Broadleaf Deciduous Boreal Shrub         | 12      | BDDBS   |
| C3 Arctic Grass                          | 13      | C3A     |
| C3 Non-Arctic Grass                      | 14      | C3G     |
| C4 Grass                                 | 15      | C4      |
| C3 Crop                                  | 16      | C3C     |
| C3 Irrigated                             | 17      | C3I     |
| Corn                                     | 18      | CO      |
| Ignated corn                             | 19      | CI      |
| Spring Temperate Cereal                  | 20      | STC     |
| Spring Temperate Cereal Irrigated        | 21      | STCI    |
| Winter Temperate Cereal                  | 22      | WTC     |
| Irrigated Winter Temperate Cereal        | 23      | WTCI    |
| Soybean                                  | 24      | SO      |
| Irrigated Soybean                        | 25      | SOI     |

Abbreviations: PFT, Plant Functional Types.
month of the timeseries over the last 50 years of the simulations. This approach to generating anomalies produces a similar result as the more common approach of subtracting the mean seasonal climatology from each year but is used here because it also removes any low frequency trends in the data. Because some grid cells show sustained trends in GPP, failing to detrend can lead to artificially high estimates of short term legacy.

2.2. Satellite and Meteorological Data

To provide a benchmark to compare the modeled estimates of legacy against, we used the 0.5° by 0.5° satellite derived estimates of GPP from Joiner et al. (2018). This product provides monthly estimates of GPP from 2001 to 2020 using a combination of SIF and MODIS reflectance along with a light use efficiency model. The product was calibrated against data from a network of global Eddy Covariance sites but the estimates do not rely on a model that is forced with meteorological data which means we can treat the estimates of legacy from the satellite data as independent from those derived from the E3SM simulations (which were forced by historical data). The primary goals here were not to use the satellite data to critique the ESM estimates of legacy but rather to ensure the simulations produced realistic estimates of legacy and therefore is a useful tool for developing hypotheses on the role of root dynamics in ecosystem legacy. We note that differences between model and satellite-estimated legacy may not be strictly from a bias in the model because the satellite estimates have their own limitations due both to a lack of sensitivity in some ecosystems and that the satellite timeseries is relatively short (20 years) with respect to properly characterizing typical responses to perturbations that are uncommon by definition. Nonetheless, we also use the satellite data to assess the broader question of whether dynamic roots improve the representation of legacy compared to simulations with fixed roots and therefore should be utilized in future studies on model simulated ecosystem responses to stress.

In addition to the satellite GPP data, we also utilize a merged satellite and meteorological aridity index (AI) product—defined as the ratio of precipitation to potential evapotranspiration—to classify the average degree of water stress at each grid cell. The AI provides a holistic perspective on water stress because it accounts for both the effects of precipitation and evaporative demand on water availability (Arora, 2002). For this study, we use the globally-derived estimates of AI from Trabucco and Zomer (2018) that combine weather data from meteorological stations (wind, precipitation, temperature, and humidity) and radiation data from MODIS to estimate both precipitation and potential evaporation from the Penman-Monteith equation. Locations with an AI greater than 1 are not chronically water-limited while sites with values less than 1 have a higher water demand than is available from precipitation. Using AI, sites can be classified along a super-arid to super-humid gradient based on their unitless AI value.

2.3. Classification of Legacy

There exists a wide range of approaches to classify legacy though they all depend fundamentally on defining how long it takes for the system to return to an unperturbed state following a forcing event (Kannenberg et al., 2019; Liu et al., 2019; Monger et al., 2015; Ogle et al., 2015). Definitions of legacy can be framed in terms of strength (i.e., how sensitive the system is to the previous state) or in terms of timescale (i.e., the length of time it takes for the system to return). For example, a system could display an initially rapid recovery but require an extended period to fully return to pre-disturbance state leading to a long legacy timescale but weak effect. For individual sites, definitions of legacy and the statistical choice can be more tailored to the local dynamics but for global analyses generalized and transferable approaches are required (Kolus et al., 2019). We opted here for two relatively simple definitions of legacy that target the strength, rather than the timescale, and can be easily implemented across ecosystems. For the first definition of legacy, we integrate the partial autocorrelation function of de-seasonalized GPP anomalies across a 1–12 months lag. This is a unitless value that captures the strength of persistence in GPP anomalies. While it is a robust metric that is widely used in many timeseries applications, it has a few critical limitations that need to be addressed in the context of this study. First, the autocorrelation function does not distinguish between endogenous (internal ecosystem processes) and exogenous (externally-forced) sources of legacy. For example, a region where precipitation anomalies persist across months would likely have a high degree of apparent GPP legacy that is not related to an ecosystem dynamic. In other words, some component of legacy defined using this approach is just the expected response of the ecosystem to an instantaneous forcing that happens to have intrinsic autocorrelation. This limitation is acceptable here because we compare simulations with and without dynamic roots that are both forced with the same climate. Therefore, differences in the
autocorrelation of GPP or Transpiration between the control and experimental simulations must arise not from the climate forcing but from the ecosystem or soil processes that are altered by the presence of dynamic roots.

The second limitation of the autocorrelation approach is that it does not distinguish between the strength of legacy associated with positive versus negative perturbations. We address this issue by utilizing a second approach to defining legacy based on identifying positive and negative GPP “events” in the timeseries’ from each grid cell and estimating the persistence of these events. In contrast to some previous studies, we did not look for droughts or pluvial events defined by climate anomalies but instead classify GPP events or perturbations as those when GPP exceeded ±1 standard deviation from the mean detrended monthly state at that grid cell. This allows us to quantify legacy across ecosystems that may have distinct limiting factors for GPP. To do this, we normalized GPP anomalies using a z-score and identified each month when the normalized GPP anomaly exceeded ±1 while also meeting the conditions that an anomaly of this size had not occurred in the previous 3 months (to avoid double counting events). We then took the sum of the normalized GPP anomalies for the 12 months following the initial event. The sum provides an integrated value for the strength of the legacy while not, per se, identifying the length of time the recovery took. The end product of these analyses are two additional definitions of legacy for each grid cell based on the direction and strength of GPP anomaly persistence following both positive and negative events.

### 2.4. Classifying the Response of Root Profiles to Perturbation

Monthly root profiles for each depth (n = 15) were converted to anomalies by detrending each month and depth of the relative fine root fraction. As with the approach to generating GPP anomalies, the detrending of the root fraction removed any seasonal cycles or long term trends in the root fraction for each depth. Using the technique to identify GPP events (either positive or negative) as described in Section 2c, we capture the root profile anomalies for the 12 months following these events. We then average the root profile anomalies (with dimensions of 12 months by 15 depths) for all GPP events within each grid cell to generate an average root profile response to positive and negative GPP events. The end product is a matrix of root profile anomalies for positive and negative GPP events for every grid cell. To identify canonical patterns in how root profiles respond to positive and negative GPP events, we utilize a k-medoid clustering algorithm to organize the global data set into a finite number of dominant root response patterns. The goal of the clustering algorithm is to divide the root profile anomalies into a pre-defined number of clusters that minimizes the sum of distances between each root profile within a cluster while maximizing the difference between clusters. Unlike k-means clustering, where the centroid of a cluster is the average of all members of the cluster, the centroid of each cluster in the k-medoid algorithm is a member of the cluster (i.e., a medoid). The clustering uses an iterative approach where first an initialization procedure identifies possible centroids and then organizes all the root profiles into the the most similar medoid—minimizing distance between each member of the cluster and the centroid. This procedure is then repeated with a new set of medoids to see if a better solution is reached, that is, where distance between clusters is increased and distance between members of the same cluster is decreased. This iterative process repeats with a new selection of medoids until no further gains are achieved. To implement this algorithm two things need to be assigned a priori: (a) the number of clusters and (b) the metric to estimate distance. We chose 4 clusters, which led to some redundancy among cluster shapes but also effectively captured key structures without having too many clusters with small populations of uncommon root responses. The distance metric was based on the Minkowski method (Kaufman & Rousseeuw, 2009), though other methods such as City Block yielded similar results.

### 2.5. Quantifying the Role of Root Dynamics in Legacy

After identifying the dominant root profiles in response to perturbation (Section 2d), we then assess whether a particular root response increases ecosystem legacy through a comparison between legacy strength in simulations with or without dynamic roots which we refer to as Δ Legacy. For example, a particular root structure that was classified as Cluster 1 might emerge in 1,000 grid cells. The legacy strength in these 1,000 grid cells forms a distribution that can be subtracted from the legacy strength of these same grid cells from the simulation that does not include dynamic roots (i.e., the Control). This process is repeated for each of the root profile clusters to assess whether particularly root responses to perturbation have the effect of adding or subtracting ecosystem legacy. We conclude the analysis by assessing whether the strongest increases or decreases in legacy associated with a particular root response can be tied to specific PFTs or climate.
3. Results

Our analysis of legacy from satellite GPP retrievals reveals a broad range of legacy strength that varies significantly by region and PFT (Figure 1). The highest levels of legacy emerge in the broadleaf deciduous temperate shrubs (BDTS) that dominate Australia, southern South America and regions of northern Mexico and the southwestern US. Following this, the broadleaf deciduous and evergreen forests prevalent in the Amazon and the maritime continent (BDTrT and BETrT) dominate the other continuous areas with high legacy. While other

Figure 1. Average legacy strength derived from the FluxSat global Gross Primary Production (GPP) product (Joiner et al., 2018). (a and b) Legacy derived from the autocorrelation of GPP anomalies in both map view and binned by Plant Functional Types (PFTs) (c and d) Legacy derived as the sum of GPP anomalies following positive GPP events in both map view and binned by PFTs. (e and f) Legacy derived as the sum of GPP anomalies following negative GPP events in both map view and binned by PFTs (Table 1). The dot in the violin plots are the median and the lines capture the 25th through 75th percentile.
regions such as Alaska, coastal regions of the western US, SE Asia, western India and the Iberian Peninsula also show regionally high legacy, the majority of PFTs show similar median values and ranges in terms of legacy strength. The spatial pattern in legacy that emerges based on the autocorrelation definition (Figures 1a and 1b), shares similarities to the legacy defined through the response of GPP anomalies to positive and negative events (Figures 1c–1f). Part of the spatial pattern in legacy strength likely reflects the length of the growing season where regions with longer growing seasons are more prone to persistent GPP anomalies whereas the boreal systems are productive for too short a period of the year to demonstrate significant intra-annual legacy. However, the range of legacy within almost all the PFTs is large, indicating the value cannot simply be explained as a function of PFT or latitude.

The control simulations without dynamic roots broadly reproduced the global patterns of legacy derived from satellite data and capture how the strength of autocorrelation varies between the PFTs (Figures 2a and 2b). The simulations do, however, typically overestimate autocorrelation as indicated by the fact that the range of data for most PFTs fall above the 1:1 line relative to satellite data. This high bias is clearly apparent in tropical Africa, parts of the Amazon and across the southern US (Figures 2a and 2b). In contrast, the simulations underestimate autocorrelation in Australia and the maritime continent. From the standpoint of providing a benchmark, the comparison with the satellite data indicates that the control simulation generates broadly realistic results across PFTs. Another critical benchmark for the control simulations is whether it produces realistic recovery rates from positive and negative GPP events. We consider this by comparing the covariation between negative and positive legacy strength for the different PFTs (Figure 3). The satellite data show a strong correlation between the legacy strength associated with positive and negative events but indicate that legacy from positive events is modestly stronger than that for negative events (Figure 3a). In other words, positive GPP anomalies are more persistent than negative ones. This effect is less well produced by the control simulations, which show less legacy associated with both positive and negative events and that the magnitude of positive and negative legacy are similar (Figure 3b). Despite these issues, the range of negative and positive legacy in the simulations falls within the range defined by the satellite benchmark.

As noted, one of the limitations to the statistical approach to define legacy here is that it does not isolate endogenous or biotic sources of persistence (e.g., pests or competition) from exogenous sources (i.e., climate or weather). While this issue does not affect our conclusions on how dynamic roots influence legacy, it is still valuable to highlight whether and how exogenous forcing influences global patterns of legacy. To assess this, we compared the legacy in GPP with temperature and precipitation autocorrelation from the same grid cells (Figure 4). The persistence in temperature and precipitation anomalies were de-trended and de-seasonalized and defined using the same unitless metrics as GPP (Section 2c) and therefore this “weather legacy” can be compared directly
against GPP legacy. The results from this analysis show how the majority of PFTs fall along a line that defines a predictable response of GPP to intra-annual persistence in weather conditions (Figures 4a and 4c). However, we also note that this linear relationship starts to fail for some of the PFTs that displayed the highest legacy including the broadleaf evergreen tropical and temperate forests and the tropical deciduous ecosystems. The former displays less legacy than expected (i.e., a weaker coupling to weather) whereas the latter shows higher GPP legacy than expected from temperature and precipitation autocorrelation. The same tight coupling of GPP to temperature and precipitation legacy is also present in the control simulations (Figures 4b and 4d). However, relative to the satellite data, the simulations show an even tighter coupling to temperature and precipitation. For example, whereas satellite data from the broadleaf evergreen ecosystems seem to deviate from the other PFTs, data from this PFT in the simulations falls more clearly along a continuum with the other ecosystems. Similarly, the broadleaf deciduous forests, do not display the distinct behavior seen in the satellite retrievals.

We focus hereafter on how legacy varies between simulations that include and exclude dynamic roots to illustrate the way root foraging modulates legacy. The presence of dynamic roots leads to an increase in legacy across the Amazon, Congo, southern Africa, northern Australia while root foraging decreases legacy in savannahs to the north and south of the Congo basin, the dry subtropical boreal forests of South America, the midwestern US, northern Europe, SE Asia and the temperate broadleaf forests in northern Europe (Figures 5a and 5b). However, most of the PFTs show no systematic change in legacy with the addition of dynamic roots but do show a large range of responses between grid cells that share the same PFT. In other words, dynamic roots did not systematically alter global patterns of legacy in one direction but the wide range of values within each PFT show that the local impact of dynamic roots on legacy strength are significant. In addition to classifying the changes of legacy by PFT, we also assessed how water stress, defined using AI, modulated the role of dynamic roots in affecting legacy (Figures 5c and 5d). This analysis shows how both wet (humid to super humid) and the most chronically water-stressed (super arid) sites displayed a shift toward increased legacy with the inclusion of dynamic roots. In contrast, the semi-arid sites showed a shift toward a decrease in legacy though variability around the mean response was large.

While we recognize there are limitations to use the satellite data to quantify the skill of the legacy strength in the simulations, we did assess whether the addition of dynamic roots systematically improved estimates of legacy. To do this, we calculated the percentage improvement in simulated legacy relative to the satellite data when dynamic roots were enabled versus when they were disabled (Figure 6). The changes in legacy associated with dynamic roots did not systematically improve nor deteriorate the quality of the simulations globally. Most PFTs

Figure 3. Relationship between scale of legacy response to positive and negative Gross Primary Production (GPP) events based on (a) satellite data (Figure 1) and (b) control simulations without dynamic roots. The global data was broken up into each of the Plant Functional Types (PFTs) in Table 1 and the range of data from each PFT is indicated by the ellipses which are color coded based on broader PFT classifications as indicated in figure. A few PFTs that are discussed in the text are marked. The ellipses encompass an area bounded by a chi-square value of 0.5 in each direction and should be taken as a relative indicator of the range of data for each PFT. The raw data used to generate these data are provided in Figure S2 in Supporting Information S1.
had percentage improvements with a median near 0 (Figure 6a). The exceptions were for corn (CO)—that showed a significant improvement—and broad evergreen shrubs (BES) ecosystems whose performance was deteriorated by the inclusion of dynamic roots. Whether the performance of a grid cell was helped or hurt by dynamic roots was not systematically organized by the AI of the site (Figure 6b). The one pattern that emerged in terms of model performance, was that sites with generally high amounts of legacy in the satellite data benefited from the addition of dynamic roots (Figure 6c).

Despite dynamic roots not systematically improving the simulated legacy, the simulations still provide useful insight into how fine root foraging can influence legacy. To understand how the presence of dynamics roots influences legacy, we assessed whether specific changes in the root profile (e.g., deepening or shallowing) enhanced legacy strength. In other words, if a negative perturbation in GPP was associated with investment in more shallow roots at one site and deeper roots at another site, we would not necessarily expect that legacy in these two sites would be affected in the same manner. As discussed in Section 2d, we used a clustering algorithm to blindly

Figure 4. Relationship between legacy strength of Gross Primary Production (GPP) versus legacy strength of temperature (a and b) and precipitation (c and d). The lines indicate a 1:1 relationship where the legacy of GPP is fully explained by the legacy of the climate forcing. The first column (a and c) shows results from the satellite data and the second column (b and d) shows the results from the control simulations with no dynamic roots. As with the previous figure, the global data was broken up into each of the Plant Functional Types (PFTs) in Table 1 and the range of data from each PFT is indicated by the ellipses which are color coded based on broader classifications as indicated in the figure. The ellipses encompass an area bounded by a chi-square value of 0.5 in each direction and should be taken as a relative indicator of the range of data for each PFT. The raw data used to generate these data are provided in Figure S3 in Supporting Information S1.
classify root structural responses, allowing us to organize grid cells with similar root responses to perturbation. As an example result, Figure 7 shows the average root response to perturbation for all grid cells from one of the clusters. As expected, following a declining GPP event, there was a relative loss of fine roots above 10 cm and an increase between 10 and 100 cm, with the change centered around 20 cm (Figure 7a). The redistribution of roots relaxes over time and the system returns to its mean root profile state after ∼10 months. An almost perfectly complementary pattern emerges following positive GPP events, where a relative accumulation of roots above 10 cm was supported by a relative loss of roots from 10 to 150 cm, centered around 80 cm (Figure 7b). As with the negative GPP response, this pattern decays and the root profile returns to its mean state by ∼10 months.

The root distribution patterns shown in Figure 7 provide an example of one of the four defined root responses to perturbation. As expected, this pattern was the most widespread global response to a negative GPP event, that is, where ecosystems gain deep roots at the expense of shallow roots (Figures 8a and 8b). Variants of this pattern are captured by two separate clusters with one being extremely widespread and associated with modest changes in root distributions (“Cluster 1” Figure 8a) and the second variation being associated with a response that was less prevalent but with a significantly larger redistribution of roots (“Cluster 2” Figure 8b). The canonical pattern that was captured by the two clusters spans sites across the globe and examples of this root response to perturbation emerged in every PFT (Figures 8a and 8b). Complementary versions of the two variations of this pattern emerged in response to positive GPP events where shallow roots increase at the expense of deeper roots (Figures 9a and 9b). As with the response to negative GPP events, the pattern of shallower roots was decomposed into two variants with one being extremely widespread but only a modest change (“Cluster 1” Figure 9a) and the other being less widespread but a more dramatic redistribution of roots (“Cluster 1” Figure 9b). The prevalence of these particular patterns reflects that the GPP events were mostly driven by changes in water availability that drove
water foraging to deeper horizons following negative GPP events and foraging for shallow water and nutrients during periods of water abundance (Drewniak, 2019; Lu et al., 2019).

Although there was a clear global dominance of the root patterns captured by Clusters 1 and 2, there were also root profile changes in response to perturbation that were significantly different in structure. For example, the grid cells that fell into Clusters 3 and 4 capture populations of sites where the model predicts a loss of deeper roots (60–100 cm) following negative GPP events and gain deeper roots following positive GPP events (Figures 8c, 8d, 9c, and 9d). The grid cells associated with Cluster 3 display a bimodal response pattern, where the loss of deep roots following negative events was also associated with marginal root loss near the surface. The number of grid cells displaying this behavior was significantly smaller and were generally found in arctic and alpine sites though scattered examples of this response can be seen elsewhere including in parts of the Sahel, Australia, and SW US.
Despite the fact that the population of sites within each cluster are defined by their similar root response to perturbation, the change in legacy (i.e., Δ Legacy) associated with sites from each cluster did not generally produce a uniform change in legacy strength or even direction. For example, the 1,000's of grid cells that fall into cluster 1, produced a wide response in legacy relative to the control. This is to say that, on average, the deepening of roots associated with water stress does not systematically increase nor decrease legacy. To gain insight into why the same root response could generate either an increase or decrease in legacy, we isolate the grid cells from within the population that display either the largest increases (90th percentile) or largest decreases (tenth percentile) in legacy relative to the control simulations and assess the conditions that define these grid cells (the analysis is illustrated in Figure 10). This analysis shows that the mean level of water stress at a site modulates whether foraging for water acts to increase or decrease ecosystem legacy (Figure 11). In the case of Cluster 1, the deeper roots following negative GPP events, had the effect of adding legacy for the non water-limited sites (i.e., AI ≥ 1) but decreased legacy at the semi-arid sites (i.e., AI ≥ 0.3 and AI ≤ 1). For Cluster 2, we also see that the deeper roots following stress events leads to a decrease in legacy at the semi-arid sites but, surprisingly, an increase in legacy at the super-arid sites.

Although the canonical pattern of deepening roots following stress events does not produce a singular type of effect on legacy, the more complex bimodal pattern associated with Cluster 3, produces a nearly universal increase in legacy for all sites that displayed this dynamic (Figure 11a). In other words, in locations where both deep and shallow roots are lost in exchange for investment in roots at intermediate depths, negative GPP anomalies were persistently extended. This occurred even though this root pattern tended to only emerge at semi-arid sites
sites which otherwise generally displayed a loss of legacy strength when dynamic roots were enabled in the model. Although the root pattern defined by Cluster 3 shares some similarity to the pattern defined by Cluster 4, the sites associated with the latter did not generate a uniform increase nor decrease in legacy. These sites that developed a shallower root profile in response to negative events follow a pattern where drier sites (all sites with $AI \leq 1$) show a decrease in legacy whereas the wetter sites show added legacy—similar to the effect of water stress on legacy as seen in sites that fell into the Cluster 1 pattern.

Although the analysis thus far has largely focused on legacy of GPP in response to perturbation, we also compare the legacy response of transpiration to dynamic roots. Legacy in transpiration is key to predicting the recovery timescale for energy partitioning (latent vs. sensible heat) and surface boundary layer coupling. On the one hand,
we assume that because transpiration and GPP are strongly coupled legacy of GPP would be mirrored by changes in the legacy of transpiration. However, it was unclear whether different root responses to GPP events (Figures 8 and 9), might modulate the relative time scales at which GPP and transpiration recover. Our comparison between the effect of dynamic roots on GPP and transpiration legacy shows there is an extremely strong similarity between the two (Figure 12). The nearly identical 1:1 response between GPP and transpiration legacy is manifest across all of the different patterns in the root response to perturbation (i.e., Clusters 1–4), indicating that all the discussion thus far on the role of roots in affecting GPP legacy can be applied when discussing transpiration legacy. Sites where dynamic roots extend GPP recovery times are also sites that show proportionately similar changes in transpiration recovery. The one exception emerged within the Broadleaf Evergreen Tropical forests (Figures 12a and 12e) (Sakschewski et al., 2021). At these sites, changes in transpiration and GPP legacy are largely decoupled and the effect of roots on GPP legacy proved modest relative to the effect of root foraging on transpiration legacy.

4. Discussion

Existing studies that have focused on ecosystem legacy have found that ESMs tend to underestimate the strength and timescale that ecosystems are affected by antecedent conditions (Kolus et al., 2019). This has been attributed to missing factors that carry previous conditions forward. Building from this work, we used a pair of general...
metrics to define legacy in GPP from both simulations and satellite data (Figures 1 and 2). We applied this analysis across all terrestrial biomes and found that the E3SM model—a state-of-the-art ESM—generally underestimated the legacy associated with both negative and positive GPP events (Figure 3). Furthermore, the simulations did not reproduce the slightly asymmetric behavior seen in the satellite data where legacy associated with positive events was stronger than from negative events (Jiang et al., 2019). This finding is somewhat novel as previous studies have tended to focus more on recovery from drought (W. R. Anderegg et al., 2015; Gonzalez-Valencia et al., 2014; Kannenberg et al., 2020) but is not surprising in that the complex ecosystem dynamics that drive legacy would optimally respond to shorten effects from stress and extend effects of positive conditions. In addition, the simulations tended to overestimate autocorrelation in GPP anomalies which we attribute to the fact that modeled estimates of productivity tend to be too tightly coupled to weather forcing which, naturally exhibits a relatively high level of persistence in most regions (Figure 4). These results largely met our a priori expectations and highlighted, as noted above, the presence of missing endogenous factors in ESMs that carry ecosystem memory forward.

Of the potential drivers that might explain the missing sources of legacy, we have focused here on the specific role that changes in the depth distribution of fine roots has on GPP and transpiration memory following perturbation. In reality, the potential role of roots on legacy is far more complex than simply a dynamic depth profile and includes changes in root morphology, hydraulic redistribution, rarified deep roots, root turnover, soil biogeochemical cycles and microbial structure. However, we use this particular aspect of root dynamics to illustrate the potentially critical and under-represented role for roots in developing more realistic legacy structure in ESMs. We intentionally amplified these effects both by increasing apparent water stress factors to drive more water foraging in the roots and extending the turnover time of fine roots so that changes in root structure persist longer. The goal was therefore less about offering a prescriptive set of model parameters to improve legacy but simply test the direction and potential sensitivity of legacy to ecosystems when roots are allowed to forage. We hypothesized that dynamic roots could force an ecosystem to be poorly conditioned when the initial stressor was relieved, leading to subsequent declines in productivity and consequently an extended memory of a perturbation such as drought. Alternatively, the adjusted root profiles could prove to better condition the system for future limitation if the cause of the initial stress (e.g., moisture stress) continued to be a persistent limiting factor. In this case, dynamic roots would reduce the legacy of the initial stressor. This hypothesized role of roots assumes roots enter a perturbation in an optimal state however, we note that due to life history, long lived species may have root profiles that are

Figure 12. The relationship between the change in ecosystem legacy (i.e., \( \Delta \text{Legacy} \)) associated with dynamic roots for GPP (x-axis) and transpiration (y-axis). The top row (A–D) shows the GPP versus transpiration legacy effects for each of the 4 clusters discussed in Figure 8 associated with positive GPP events. The bottom row shows (E–H) shows the changes in GPP versus transpiration legacy for each of the 4 clusters discussed in Figure 9 associated with negative GPP events. In panels A and E, the points associated with Broadleaf Evergreen Tropical Forests are shown in pink because data from these grid cells had a clearly distinct response. The red line in all plots show the 1:1 line for reference. The black lines in Panels A and E are used to show the fit to the data from Broadleaf Evergreen Tropical Forests. The coloring of the data is used to indicate relative density of points.
pre-conditioned for extreme stress events. This would further determine the sensitivity of roots profiles to stress. Our results from the simulations show that depending on the background water stress of a site, both enhanced and reduced legacy emerge as a result of root foraging (Figures 5 and 11).

One of the first challenges to understanding the role of roots in driving legacy, was the question of how roots in fact respond to perturbation. Overwhelmingly, the simulations suggest that across all climatic and PFT groups—irrespective of the initialized root profile—there was a deepening of roots following negative GPP events (Figure 7). This confirms that within the simulations, water stress was the dominant driver of negative GPP events and that foraging for water systematically leads to deeper root profiles (Canadell et al., 1996; Doughty et al., 2014; Drewniak, 2019; Grossiord et al., 2017). The modeled changes in root distribution were generally modest which reflects that foraging deeper comes at hydraulic costs and the expense of nutrient access and so a radical redistribution of roots would be highly detrimental (Dybzinski et al., 2011). The changes in root distribution tended to shift around an inflection at 10 cm (loss of roots above this depth) and the altered profile tended to fade away within 8–10 months (Figure 7). Both the direction and magnitude of the simulated changes in root profiles were broadly consistent with site level observations (Joslin et al., 2000; Schenk & Jackson, 2002). However, because the model imposes limits on maximum rooting depth, the simulations miss the complex role that rarified deep roots play in providing a buffer from water stress while allowing for shallow roots to continue to access nutrients and take advantage of small rainstorms (Germon et al., 2020; Jarvis, 2011). Nonetheless, the effect on legacy that comes from ecosystems deepening their roots in response to negative stressors, can be explained through a set of simple conceptual models that we describe below.

In semi-arid ecosystems that are typically water stressed, the deeper root profile after negative stress events remains favorable and allows these systems—whether they are prescribed as forests, shrubs or grasslands—to recover more quickly. However, in these same systems the shallowing of the root system following positive events tends to leave the vegetation vulnerable to subsequent water stress—which is the norm for systems with AI values of less than 1—and thus the systems cannot maintain positive GPP anomalies following the initial event. In other words, dynamic roots diminish the legacy of both positive and negative events in semi-arid systems. A similar effect emerged in a forest irrigation experiment described by Zweifel et al. (2020), where following multiple years of irrigating, a semi-arid forest invested more heavily in root systems. When the irrigation experiment ended, the forest under-performed the control during dry periods of the year. In other words, this artificially-imposed pluvial event led the system to be poorly conditioned for the more normally water stressed state of the ecosystem.

The modeled behavior in semi-arid ecosystems was inverted in wetter systems where the deeper root profile following negative events led to sustained reductions in productivity because the deeper root profile reduced access to nutrients while the increased access to deeper soil water was not favorable. These systems also shallowed their roots during positive GPP events, which sustained the positive productivity anomalies because of the increased access to nutrients at minimal expense to the lost access to deeper water pools. In other words, wetter regions showed an increased legacy in response to root foraging (Figure 11). The more surprising result was observed in the hyper-arid sites which, unlike the semi-arid sites, showed enhanced legacy. We hypothesize that this dynamic arose because the deeper root profiles following negative GPP events, led to reduced access to surface soil moisture pools from small precipitation events and therefore extended legacy from stress events. Similarly, the shallower roots after positive events enhanced access to these surface water pools and thus extended the positive GPP anomaly. The ability for the severely water-stressed ecosystem to take advantage of these small rain events has also been observed outside of simulations (Ritter et al., 2020) and so whereas the deeper root system will expedite recovery in semi-arid systems, the driest systems are penalized by lost access to near surface soil water (O. E. Sala et al., 2012). The actual dynamics in these ecosystems are likely to be strongly affected by the development of dimorphic root profiles with some extremely deep roots (Dawson & Pate, 1996). The inability of the model to reproduce very deep roots means that the modeled effect of roots on legacy in these very dry systems is somewhat artificial. Despite the more counter-intuitive results in the hyper-arid systems, the broad pattern of deepening roots from negative events and shallowing roots from positive events leads to effects on legacy that can be readily predicted based on whether a system is generally water stressed or not (O. E. Sala et al., 2012).

Although most of the ecosystems in these simulations behaved as water-limited and thus foraged for deep water, a smaller subset of grid cells had strongly competing water and nutrient limitations which resulted in the root response captured by Cluster 3 where both shallow and deep roots were lost—the opposite of root dimorphism...
These grid cells displayed a response to stress showing loss of surface roots (≤2 cm) and deep roots (≥80 cm) while roots accumulated in the more intermediate depths. This pattern emerged as a consequence of a trade-off between optimizing water access by allocating fine roots between 10 and 80 cm at the expense of deeper roots that can supply only limited nutrients. This pattern only emerged in sites where AI was less than 1 but, interestingly, lead to an increase in legacy in 90% of the grid cells that displayed this root allocation behavior. This dynamic is notable because it is opposite to the more common reduction in legacy at semi-arid sites discussed above. We hypothesize that following negative GPP events, the loss of both the shallow and deep root pools, sustains the negative GPP anomalies by extending both water and nutrient limitations. On the other hand, the added shallow and deeper roots after positive GPP events, can sustain the positive anomalies by meeting the demands of both the persistent water and nutrient limitations. This pattern does not emerge in sites where the driver of stress is singularly water. We highlight this case because it shows how the competing demands for nutrient and water foraging led to a more complex pattern and one that exclusively enhances legacy, which, as noted, is a chronic issue in ESMs.

The simulations suggest that the effect of dynamic roots on GPP legacy can be largely explained through a consideration of background water stress and whether that co-exists with nutrient limitations. We also show that these changes in GPP legacy are equivalently mirrored by changes in transpiration legacy. The response of transpiration legacy relative to GPP was largely unaffected by the structure of the root response to stress such that, for example, the common pattern where semi-arid systems deepen their roots and expedite GPP recovery holds true for transpiration as well. One implication of this is that the more rapid recovery of transpiration as a result of dynamic roots means that the canopy will experience less sustained reductions in latent heat (or increases in sensible heat) (Yunusa et al., 2015). The presence of dynamic roots in semi-arid systems thus might reduce the effect of land-atmosphere exchange in sustaining surface-boundary layer feedbacks associated with drought. In addition, the more rapid recovery of transpiration would also mean that vegetation water demands would recover at the expense of, for example, enhanced runoff following a negative GPP event. This latter point goes beyond the direct topic of this study but is interesting to consider the effect that dynamic roots have on recovery from the perspective of interannual variance in streamflow generation (Brooks et al., 2021).

The strong coupling between GPP and transpiration is, for the most part, expected though we also note that the two fluxes can become more loosely coupled during periods of limited water stress where radiation or nutrients might limit GPP but transpiration can remain high (Berkelhammer et al., 2020). Indeed, we note one key deviation between the response of GPP and transpiration legacy to dynamic roots emerged in the tropical broadleaf evergreen systems that showed a much wider range in the response of transpiration legacy to dynamic roots relative to GPP. These are sites that were already characterized by relatively high levels of GPP legacy (Figure 1) and this effect is thus amplified in terms of how transpiration anomalies persist. Evaporative demand is particularly high year round at these sites and slight modifications of root profiles to optimize water access, has an amplified and sustained effect on water use. While this behavior is unique to this PFT, we note that this ecosystem is widespread and transpiration from this region is critical not only to supporting regional rainfall patterns (Wright et al., 2017) but also to global humidity budgets (Worden et al., 2007). Consequently, extending the transpiration anomalies following a drought or pluvial event in the tropical evergreen forests have impacts that may be widespread. However, simulations with a coupled land and atmosphere would be needed to quantify this effect.

The discussion thus far has focused on the way the inclusion of dynamic roots influences simulated legacy with minimal emphasis on whether this actually improves model performance. An obvious question is thus whether dynamic roots should be enabled in ESMs that already include this module (such as E3SM) or developed for models that do not. Drewniak (2019) has already shown that dynamic roots yielded only modest improvements in simulated global patterns of GPP and ET. From this perspective, there is minimal incentive to utilize dynamic roots as the benefits in terms of global patterns of carbon and water exchange likely do not warrant the added complexity and/or computational costs it could add (Moore, 2022). The reason root foraging has only a small impact on mean carbon and water fluxes is that as soil layers dry out from transpiration demands, soil moisture is redistributed to drier soil layers through the root system or soil layers with low density of roots can contribute more to total water use—that is, compensatory flow (Jarvis, 2011). The model is thus able to compensate for the role of root foraging for water with a static root profile. However, the perspective gained from comparing mean global patterns GPP or ET in the model versus observations does not confront the role that adding this process has in affecting transient components of the carbon and water flux. In other words, mean values might
be unaffected by adding a process such as dynamic roots but the timescale over which a system responds and recovers from stress might be modified by this process. For example, the timescale that water can be redistributed by roots is extremely fast relative to the timescale needed to significantly re-allocate fine roots to wetter soil layers. Indeed, in numerous ways through this paper we have shown that adding dynamic roots influences the way ecosystems across the globe respond to perturbation despite only minimally influencing total productivity. Here, again, adding dynamics roots did not substantively improve (relative to satellite derived metrics of GPP) global patterns of legacy writ large. However, grid cells with large a priori amounts of legacy did demonstrate improved performance. Site level studies with better observations for benchmarking (i.e., eddy covariance towers) are needed to diagnose why the simulations showed improved performance at these grid cells but it could reflect increased coupling of vegetation to weather by allowing roots to quickly access available water. That said, from the standpoint of whether the inclusion of this process improves performance, the results are ambiguous. In terms of other standards to consider when adding a new ecological process to an ESM as posed by Moore (2022) and Kyker-Snowman et al. (2021): (a) dynamic roots do operate within the pre-existing structure of the model; (b) there is both sufficient theory and (c) a mathematical framework for developing dynamic roots; and (d) a growing community is developing observational datasets (such as NEON's network of Minirhizotrons) that would aid in better development and testing of this process. The benefits of adding dynamic roots to an ESM might become more apparent when utilized in parallel with other processes being developed for ESMs. For example, schemes with more complex vegetation hydraulics could be run with and without dynamic roots to understand the costs of drawing water from different layers (Kennedy et al., 2019) against the cost of redistributing roots or the how hydraulic redistribution limits foraging or facilitates more foraging. Furthermore, the study did not address the potential biogeochemical implications of dynamic roots which could enhance performance in carbon soil cycling that might not be evident from benchmarking against satellite reflectance data as done here.

5. Conclusion

Earth System Models lack the ecosystem dynamics needed to capture the full spectrum of legacy effects associated with climate events such as drought. We have shown that when roots are allowed to forage for water in an ESM, the depth distribution of roots remains altered long after the initial stressor is relieved generating belowground structural legacy similar to the way reduced woody biomass persists for years after droughts. While the changes in root distribution and persistence after the initial event are similar across ecosystems, the effect of root structural changes on the legacy of ecosystem carbon and water fluxes varies depending primarily on the water stress of the site. In semi-arid ecosystems, the outward manifestation of root structural legacy is a more rapid return to the pre-stressed state. In other words, the belowground structural legacy reduced the functional legacy. This is complementary to some field observations such as those from Kannenberg et al. (2020), who noted that GPP quickly rebounded from a drought while canopy structure remained perturbed. In contrast, wetter ecosystems were penalized for altering their root profile and show a reduced rate of recovery—that is, an increase in legacy from stress events. While the analysis displayed systematic ways that root foraging influences ecosystem legacy, the addition of dynamic roots did not enhance nor degrade the performance of the model relative to legacy estimates derived from satellite retrievals. So while the analysis did not yield a clear prescriptive rule for when or where to include dynamic roots in ESMs, the results suggest that studies focusing on recovery from disturbance or effects of extreme events should utilize dynamic roots to generate a spectrum of recovery timescales particular in the context of the compounding effects of increasingly frequent stress events (Szejner, 2020). Future work would benefit from the inclusion of more holistic root dynamic including dynamic morphology, more complexity to the turnover time of root carbon pools (Matamala et al., 2003), dynamic allocation schemes between aboveground and belowground pools Lu et al. (2019) as well as the addition of deeper root pools (Fan et al., 2017) and hydraulic redistribution. As the representation of root processes becomes more complex in ESMs and a wider observational network becomes available, testing the effects of these dynamics on ecosystem recovery will enable field-testable hypotheses about the role of belowground root structure as a source of ecosystem legacy on land surface-atmosphere coupling.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.
Data Availability Statement

All simulations utilized in this manuscript have been archived and made publicly through the ESS-DIVE data repository at https://doi.org/10.15485/1814209. The aridity index data was utilized as previously published and downloaded from: https://cgiarcsi.community/data/global-aridity-and-plant-database/. The satellite Gross Primary Production (GPP) data (i.e., FLUXSAT) was downloaded from: https://avdc.gsfc.nasa.gov/pub/tmp/FluxSat_GPP/.

Acknowledgments

The authors would like to acknowledge DOE grant DE-SC0020285 to M. Berkelhammer, M. A. Gonzalez-Meler, and B. Drewniak for supporting this work.

References

Amenu, G. G., Kumar, P., & Liang, X.-Z. (2005). Interannual variability of deep-layer hydrologic memory and mechanisms of its influence on surface energy fluxes. Journal of Climate, 18(23), 5024–5045. https://doi.org/10.1175/jcli3590.1

Anderegg, L. D., Anderegg, W. R., Abatzoglou, J., Hausladen, A. M., & Berry, J. A. (2013). Drought characteristics’ role in widespread aspen forest mortality across Colorado, USA. Global Change Biology, 19(5), 1526–1537. https://doi.org/10.1111/gcb.12146

Anderegg, W. R., Schwalm, C., Biondi, F., Camarero, J. J., Koch, G., Litvak, M., et al. (2015). Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. Science, 349(6247), 528–532. https://doi.org/10.1126/science.aab1833

Arora, V. K. (2002). The use of the aridity index to assess climate change effect on annual runoff. Journal of Hydrology, 265(1–4), 164–177. https://doi.org/10.1016/s0022-1694(02)00104-4

Belk, E. L., Markewitz, D., Rasmussen, T. C., Carvalho, E. J. M., Nepstad, D. C., & Davidson, E. A. (2007). Modeling the effects of through-fall reduction on soil water content in a Brazilian Oxisol under a moist tropical forest. Water Resources Research, 43(8). https://doi.org/10.1029/2006wr005493

Berkelhammer, M., Still, C., Ritter, F., Winnick, M., Anderson, L., Carroll, R., et al. (2020). Persistence and plasticity in conifer water-use strategies. Journal of Geophysical Research: Biogeosciences, 125(2), e2019JG008485. https://doi.org/10.1029/2019JG008485

Bloom, A. J., Chapin III, F. S., & Mooney, H. A. (1985). Resource limitation in plants—an economic analogy. Annual Review of Ecology and Systematics, 16(1), 363–392. https://doi.org/10.1146/annurev.es.16.110185.002051

Bouda, M., & Saiers, J. E. (2017). Dynamic effects of root system architecture improve root water uptake in 1-d process-based soil-root hydrodynamics. Advances in Water Resources, 110, 319–334. https://doi.org/10.1016/j.advwatres.2017.10.018

Brooks, P. D., Gelderloos, A., Wolf, M. A., Jamison, L. R., Strong, C., Solomon, D. K., et al. (2021). Groundwater-mediated memory of past climate controls water yield in snowmelt-dominated catchments. Water Resources Research, 57(10), e2021WR028005. https://doi.org/10.1029/2021wr028005

Brunner, I., Pannatier, E. G., Frey, B., Rigling, A., Landolt, W., Zimmermann, S., & Dobbertin, M. (2009). Morphological and physiological responses of Scots pine fine roots to water supply in a dry climatic region in Switzerland. Tree Physiology, 29(4), 541–550. https://doi.org/10.1093/treephys/tpn046

Bunde, A., Blütschi, U., Ludescher, J., Luterbacher, J., & Von Storch, H. (2013). Is there memory in precipitation? Nature Climate Change, 3(3), 174–175. https://doi.org/10.1038/nclimate1830

Burrows, S., Maltrad, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., et al. (2020). The DOE E3SM v1.1 biogeochemistry configuration: Description and simulated ecosystem-climate responses to historical changes in forcing. Journal of Advances in Modeling Earth Systems, 12(9), e2019MS001766. https://doi.org/10.1029/2019ms001766

Campioli, M., Malhi, Y., Vicca, S., Luyssaert, S., Papale, D., Peñuelas, J., et al. (2016). Evaluating the convergence between eddy-covariance and biometric methods for assessing carbon budgets of forests. Nature Communications, 7(1), 1–12. https://doi.org/10.1038/ncomms13717

Canadell, J., Jackson, R., Ehleringer, J., Mooney, H., Sala, O., & Schulze, E.-D. (1996). Maximum rooting depth of vegetation types at the global scale. Oecologia, 108(4), 583–595. https://doi.org/10.1007/bf00292030

Dawson, T. E., & Pate, J. S. (1996). Seasonal water uptake and movement in root systems of Australian phraeatophytic plants of dimorphic root morphology: A stable isotope investigation. Oecologia, 107(1), 13–20. https://doi.org/10.1007/bf00582230

Dewar, R. C., Ludlow, A. R., & Dougherty, P. M. (1994). Environmental influences on carbon allocation in pines. Ecological Bulletins, 92–101.

Dirmeyer, P., Gao, X., & Oki, T. (2002). The response of tropical forest growth to seasonal and interannual drought. Water Resources Research, 38(11), 1771. https://doi.org/10.1029/2001wr000920

Donat, M. G., Pitman, A. J., & Angélil, O. (2018). Understanding and reducing future uncertainty in midlatitude daily heat extremes via land surface feedback constraints. Geophysical Research Letters, 45(19), 10–627. https://doi.org/10.1029/2018gl079128

Doughty, C. E., Malhi, Y., Araujo-Murakami, A., Metcalfe, D. B., Silva-Espejo, J. E., Arroyo, L., et al. (2019). Simulating dynamic roots in the energy exascale Earth system land model. Journal of Advances in Modeling Earth Systems, 11(1), 338–359. https://doi.org/10.1029/2019JG005133

Dybrynski, R., Farrier, C., Wolf, A., Reich, P. B., & Pacala, S. W. (2011). Evolutionarily stable strategy carbon allocation to foliage, wood, and fine roots in trees competing for light and nitrogen: An analytically tractable, individual-based model and quantitative comparisons to data. The American Naturalist, 177(2), 153–166. https://doi.org/10.1086/657992

El Mushri, B., Shu, S., & Jain, A. K. (2015). Implementation of a dynamic rooting depth and phenology into a land surface model: Evaluation of carbon, water, and energy fluxes in the high latitude ecosystems. Agricultural and Forest Meteorology, 211, 89–99. https://doi.org/10.1016/j.agrformet.2015.06.002

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(72), 3618–3623. https://doi.org/10.1029/2019JG005133

Fischer, E. M., Seneviratne, S. I., Lüthi, D., & Schär, C. (2007). Contribution of land-atmosphere coupling to recent European summer heat waves. Geophysical Research Letters, 34(6). https://doi.org/10.1029/2006gl029068

Flower, C. E., & Gonzalez-Meler, M. A. (2015). Responses of temperate forest productivity to insect and pathogen disturbances. Annual Review of Plant Biology, 66, 547–569. https://doi.org/10.1146/annurev-arplant-040314-115540

Galiano, L., Martínez-Vilalta, J., & Lloret, F. (2011). Carbon reserves and canopy defoliation determine the recovery of Scots pine 4 yr after a drought episode. New Phytologist, 190(3), 750–759. https://doi.org/10.1111/j.1469-8137.2010.03828.x
Ovenden, T. S., Perks, M. P., Clarke, T. K., Mencuccini, M., & Jump, A. S. (2021). Life after recovery: Increased resolution of forest resilience assessment sheds new light on post-disturbance recovery and growth dynamics. *Journal of Ecology.*

Peltier, D. M., & Ogle, K. (2020). Tree growth sensitivity to climate is temporally variable. *Ecology Letters,* 23(11), 1561–1572. https://doi.org/10.1111/ele.13575

Phillips, R. P., Ibáñez, I., D’Orangeville, L., Hanson, P. J., Ryan, M. G., & McDowell, N. G. (2016). A belowground perspective on the drought sensitivity of forests: Towards improved understanding and simulation. *Forest Ecology and Management,* 380, 309–320. https://doi.org/10.1016/j.foreco.2016.08.043

Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., & Mommer, L. (2012). Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. *New Phytologist,* 193(1), 30–50. https://doi.org/10.1111/j.1469-8137.2011.03952.x

Richardson, A. D., Carbone, M. S., Huggett, B. A., Furze, M. E., Czimczik, C. I., Walker, J. C., et al. (2015). Distribution and mixing of old and new nonstructural carbon in two temperate trees. *New Phytologist,* 206(2), 590–597. https://doi.org/10.1111/nph.13273

Ritter, F., Berkelhammer, M., & Garcia-Eidell, C. (2020). Distinct response of gross primary productivity in five terrestrial biomes to precipitation variability. *Communications Earth & Environment,* 1(1), 1–8. https://doi.org/10.1038/s43247-020-00034-1

Sakschewski, B., von Bloh, W., Driese, M., Sörensson, A. A., Rusca, R., Langerschmeer, F., et al. (2021). Variable tree rooting strategies are key for modelling the distribution, productivity and evapotranspiration of tropical evergreen forests. *Biogeosciences,* 18(13), 4091–4116. https://doi.org/10.5194/bg-18-4091-2021

Sala, O. E., Gherardi, L. A., Reichmann, L., Jobbagy, E., & Peters, D. (2012). Legacies of precipitation fluctuations on primary production: Theory and data synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1606), 3135–3144. https://doi.org/10.1098/rstb.2011.0347

Sala, O., Lauenroth, W., & Parton, W. (1992). Long-term soil water dynamics in the shortgrass steppe. *Ecology,* 73(4), 1175–1181. https://doi.org/10.2307/1940667

Schenk, H. J., & Jackson, R. B. (2002). Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. *Journal of Ecology,* 90(3), 480–494. https://doi.org/10.1046/j.1365-2745.2002.00682.x

Seddon, A. W., Macias-Fauria, M., Long, P. R., Benz, D., & Willis, K. J. (2016). Sensitivity of global terrestrial ecosystems to climate variability. *Nature,* 531(7593), 229–232. https://doi.org/10.1038/nature16986

Szejner, P., Belmecheri, S., Ehleringer, J. R., & Monson, R. K. (2020). Recent increases in drought frequency cause observed multi-year drought legacies in the tree rings of semi-arid forests. *Oecologia,* 192(1), 241–259. https://doi.org/10.1007/s00442-019-04550-6

Thorley, J. H. (1998). Modelling shoot [ratio] root relations: The only way forward? *Annals of Botany,* 81(2), 165–171. https://doi.org/10.1006/anbo.1997.0529

Thornton, P. E., & Rosenbloom, N. A. (2005). Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model. *Ecological Modelling,* 189(1–2), 25–48. https://doi.org/10.1016/j.ecolmodel.2005.04.008

Trabucco, A., & Zomer, R. J. (2018). Global aridity index and potential evapotranspiration (ET0) climate database v2. CGIAR Consortium for Spatial Information.

Trugman, A., Detto, M., Bartlett, M., Medvigy, D., Anderegg, W., Schwalm, C., et al. (2018). Tree carbon allocation explains forest drought-kill and recovery patterns. *Ecology Letters,* 21(10), 1552–1560. https://doi.org/10.1111/ele.13136

van der Molen, M. K., Dolman, A. J., Ciais, P., Eglin, T., Gobron, N., Law, B. E., et al. (2011). Drought and ecosystem carbon cycling. *Agricultural and Forest Meteorology,* 151(7), 765–773. https://doi.org/10.1016/j.agrformet.2011.01.018

Warren, J. M., Hanson, P. J., Iversen, C. M., Kumar, J., Walker, A. P., & Wullschleger, S. D. (2015). Root structural and functional dynamics in terrestrial biosphere models—evaluation and recommendations. *New Phytologist,* 205(1), 59–78. https://doi.org/10.1111/nph.13034

Worden, J., Noone, D., Bowman, K., & Bowman, K. (2007). Importance of rain evaporation and continental convection in the tropical water cycle. *Quarterly Journal of the Royal Meteorological Society,* 133(692), 528–532. https://doi.org/10.1002/qj.5508

Wright, J. S., Fu, R., Worden, J. R., Chakraborty, S., Clinton, N. E., Risi, C., et al. (2017). Rainforest-initiated wet season onset over the southern amazon. *Proceedings of the National Academy of Sciences,* 114(32), 8481–8486. https://doi.org/10.1073/pnas.1621515114

Yumusa, I. A., Eamus, D., Taylor, D., Whitley, R., Gwerni, W., Palmer, A. R., & Li, Z. (2015). Partitioning of turbulent flux reveals contrasting cooling potential for woody vegetation and grassland during heat waves. *Quarterly Journal of the Royal Meteorological Society,* 141(692), 2528–2537. https://doi.org/10.1002/qj.2539

Zeng, X. (2001). Global vegetation root distribution for land modeling. *Journal of Hydro meteorology,* 2(5), 525–530. https://doi.org/10.1175/1525-7541(2001)002<0525:grvrdf>2.0.co;2

Zweifel, R., Elzolt, S., Sterck, F., Gessler, A., Anfodillo, T., Mencuccini, M., et al. (2020). Determinants of legacy effects in pine trees–implications from an irrigation-stop experiment. *New Phytologist,* 227(4), 1081–1096. https://doi.org/10.1111/nph.16382