https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

Upscaling dryland carbon and water fluxes with artificial neural networks of optical, thermal, and microwave satellite remote sensing

Matthew P. Dannenberg[1], Mallory L. Barnes[2], William K. Smith[3], Miriam R. Johnston[1], Susan K. Meerdink[1], Xian Wang[2,3], Russell L. Scott[4], Joel A. Biederman[4]

[1]Department of Geographical and Sustainability Sciences, University of Iowa, Iowa City IA 52245, USA
[2]O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington IN 47405, USA
[3]School of Natural Resources and the Environment, University of Arizona, Tucson AZ 85721, USA
[4]Southwest Watershed Research Center, Agricultural Research Service, U.S. Department of Agriculture, Tucson AZ 85719, USA

Correspondence to: Matthew P. Dannenberg (matthew-dannenberg@uiowa.edu)

Abstract. Earth's drylands are home to more than two billion people, provide key ecosystem services, and exert a large influence on the trends and variability in Earth's carbon cycle. However, modeling dryland carbon and water fluxes with remote sensing suffers from unique challenges not typically encountered in mesic systems, particularly in capturing soil moisture stress. Here, we develop and evaluate an approach for joint modeling of dryland gross primary production (GPP), net ecosystem exchange (NEE), and evapotranspiration (ET) in the western United States (U.S.) using a suite of AmeriFlux eddy covariance sites spanning major functional types and aridity regimes. We use artificial neural networks (ANNs) to predict dryland ecosystem fluxes by fusing optical vegetation indices, multitemporal thermal observations, and microwave soil moisture/temperature retrievals from the Soil Moisture Active Passive (SMAP) sensor. Our new dryland ANN (DrylANNd) carbon and water flux model explains more than 70% of monthly variance in GPP and ET, improving upon existing MODIS GPP and ET estimates at most dryland eddy covariance sites. DrylANNd predictions of NEE were considerably worse than its predictions of GPP and ET, likely because soil and plant respiratory processes are largely invisible to satellite sensors. Optical vegetation indices, particularly the normalized difference vegetation index (NDVI) and near-infrared reflectance of vegetation (NIR_v), were generally the most important variables contributing to model skill. However, daytime and nighttime land surface temperatures and SMAP soil moisture and soil temperature also contributed to model skill, with SMAP especially improving model predictions of shrubland, grassland, and savanna fluxes and land surface temperatures improving predictions in evergreen needleleaf forests. Our results show that a combination of optical vegetation indices, thermal infrared, and microwave observations can substantially improve estimates of carbon and water fluxes in drylands, potentially providing the means to better monitor vegetation function and ecosystem services in these important regions that are undergoing rapid hydroclimatic change.

1 Introduction

Earth's drylands are critically important to society yet exceptionally vulnerable to climate change. Drylands are home to more than two billion people and make up more than 40% of Earth's land surface (Reynolds et al., 2007). Primary production of dryland vegetation supports many rare and endemic species as well as extensive rangelands and croplands (Bestelmeyer et al., 2015). Dryland ecosystems are also important regulators of global trends and interannual variability of Earth's carbon cycle (Humphrey et al., 2018; Ahlström et al., 2015; Poulter et al., 2014), due both to their large spatial extent and high climate sensitivity (Biederman et al., 2016; Zhang et al., 2022). Hotter and atmospherically drier conditions associated with anthropogenic climate change will likely increase the intensity of water limitation (Cayan et al., 2010; Cook et al., 2015; Williams et al., 2020; Ault, 2020; Cook et al., 2020), possibly leading to expansion and degradation of drylands (Huang et al., 2016, 2017). There is therefore a pressing need for satellite-based monitoring of dryland carbon and water cycling at large scales.

While many remote sensing techniques were originally developed and tested in drylands (e.g., Huete, 1988; Huete & Jackson, 1987; Rouse et al., 1974), satellite-based modeling of dryland carbon and water fluxes has been a long-standing challenge. For example, early validation studies of the Moderate Resolution Imaging Spectroradiometer (MODIS) science products noted tendencies to overestimate mean dryland productivity (Heinsch et al., 2006; Turner et al., 2005, 2006a, b) and to miss important features of the seasonal cycle (Heinsch et al., 2006; Turner et al., 2006b). Recent work has also shown that both satellite models (Biederman et al., 2017; Stocker et al., 2019) and process-based models (MacBean et al., 2021) dramatically underestimate the interannual variability of dryland carbon and water fluxes, while also frequently failing to capture the "flashy" and multi-modal seasonal dynamics characteristics of dryland carbon cycling (Barnes et al., 2021). For instance, the widely used MODIS gross primary production (GPP) and evapotranspiration (ET) products drastically underestimate the variability of carbon and water fluxes of the western US, capturing only ~30% of interannual variability (Biederman et al., 2017).

Several issues make drylands uniquely difficult to monitor and model with remote sensing (Smith et al., 2019). First, ecosystem carbon and water exchange are more tightly coupled to soil moisture in drylands than in more mesic systems (Novick et al., 2016; Stocker et al., 2018), but most existing satellite-based models do not explicitly represent soil moisture stress (Song et al., 2013). Instead, light-use efficiency models often represent moisture stress using vapor pressure deficit (Running et al., 2004; Zhang et al., 2016), which is well suited for mesic regions but not drylands, where soil moisture plays a particularly important role in regulating surface conductance and carbon and water fluxes (Novick et al., 2016; Stocker et al., 2018; Dannenberg et al., 2022). The effects of soil moisture stress therefore need to be incorporated into satellite-based carbon and water models to fully capture temporal variability of dryland water limitation (Stocker et al., 2018, 2019; Smith et al., 2019).

Second, dryland plants have physiological responses to water limitation (and precipitation variability more generally) that are not necessarily captured by standard remote sensing approaches. Many dryland plants have drought adaptations that allow them to remain green even while functionally inactive under extreme moisture stress (Yan et al., 2019; Smith et al., 2019), making it difficult to resolve temporal variation in dryland plant function. Therefore, plant physiological responses to periods of moisture stress are not necessarily reflected in optical vegetation indices (VIs) (Yan et al., 2019; Wang et al., 2022; Smith et al., 2018) or satellite-based estimates of the fraction of absorbed photosynthetically active radiation by the canopy (Turner et al., 2006a). Combining different types of remotely sensed observations—such as those from microwave, thermal, and visible wavelengths—can capture complementary information about plant and ecosystem stress that is unattainable from optical VIs alone (Smith et al., 2019; Stavros et al., 2017; Guan et al., 2017). Integration of multi-source satellite remote sensing could therefore improve representation of plant physiological responses to periodic moisture stress in drylands as compared with optical VIs alone.

Third, drylands tend to be more spatially heterogeneous than many other ecosystems, consisting of complex mixtures of vegetation structural, morphological, functional, and physiological characteristics that vary over relatively short distances. These mixtures of vegetation types within moderate- to coarse-resolution imagery can contribute to significant error in GPP estimates (Turner et al., 2002; Heinsch et al., 2006). Many large-scale remote sensing-based carbon and water models assume a single vegetation type for each coarse pixel rather than representing the land surface as a continuous mixture of different cover types. High spatial heterogeneity in dryland vegetation, in combination with complex terrain in some areas, leads to diverse ecosystem seasonalities. Drylands in the western U.S. often have one or more annual growing seasons occurring in spring and/or summer (Biederman et al., 2017; Dannenberg et al., 2020), and the timing, length, and productivity of those growing seasons can vary substantially from year to year in response to ocean-atmosphere teleconnections (Dannenberg et al., 2015, 2021). Moreover, carbon and water fluxes in drylands depend on intermittent and highly variable “pulses” of precipitation that are less seasonally and spatially uniform than limiting resources (e.g., temperature and light) in more mesic or temperate ecosystems (Huxman et al., 2004; Roby et al., 2020). This combination of high spatial and temporal heterogeneity in dryland ecosystem structure and function leads to highly “unique” patterns in carbon and water fluxes, meaning that models perform poorly when used to predict fluxes at sites on which they are not trained (Haughton et al., 2018). The flux tower networks typically used to train remote sensing-based models, however, have notably low representation of dryland sites relative to their global prevalence (Smith et al., 2019). The spatial and temporal heterogeneity of limiting resource availability, “uniqueness” of dryland fluxes, and relatively sparse dryland observation networks combine to increase uncertainty in carbon and water cycling estimates from models primarily calibrated for other places and other types of ecosystems.

With new sensors, new vegetation indices, and expanded global ground networks, many of these issues are now at least partly addressable. Recent research has focused, for example, on using different combinations of remote sensing data, including integration of soil moisture (Stocker et al., 2019; Jones et al., 2017), multispectral (Barnes et al., 2021), and thermal infrared

(Sims et al., 2008; Anderson et al., 2012) observations in models ranging in complexity from purely empirical to semi-empirical or process-based. Barnes et al. (2021), for example, developed dryland-specific GPP estimates based on machine learning of meteorological reanalysis data and optical remote sensing observations and showed that models containing hydrological variables and trained specifically for drylands outperform globally-trained models at capturing seasonal to interannual variability of dryland GPP. New satellite microwave missions also allow more direct sensing of soil moisture than previously available (Song et al., 2013; Jones et al., 2017; Smith et al., 2019), which could address one of the biggest contributors to model error in dryland ecosystems: the tight coupling between plant activity and soil moisture that is not well-captured by vapor pressure deficit (Novick et al., 2016; Stocker et al., 2018; Heinsch et al., 2006) or remotely sensed greenness (Yan et al., 2019).

Here, we develop and test an approach for data-driven prediction of a full suite of carbon and water fluxes that are specifically adapted for drylands using a machine learning fusion of multispectral, thermal, and microwave remote sensing. We use an ensemble of artificial neural networks (ANNs) to jointly predict the key ecosystem carbon and water fluxes—GPP, net ecosystem exchange (NEE), and ET—at monthly, 0.05° resolution using a combination of optical VIs from MODIS, daytime and nighttime MODIS land surface temperatures (LST), soil moisture and soil temperature from the Soil Moisture Active Passive (SMAP) sensor, and subpixel fractional land cover. We develop and test the dryland ANN model (hereafter called "DrylANNd") using flux observations from 28 AmeriFlux eddy covariance towers in arid to subhumid regions of the western United States. We evaluate the ability of the model to capture monthly variability, spatial patterns, seasonality, and interannual variability of GPP, NEE, and ET; compare the model estimates to existing MODIS GPP and ET products; and assess which remotely sensed variables are most important for improving GPP, NEE, and ET estimates in drylands. Our data-driven model will provide new and improved estimates of the variability and hydroclimatic drivers of carbon and water fluxes across the western U.S., with the potential to inform and develop future global-scale estimates.

2 Materials and Methods

2.1 Study area and eddy covariance data

We developed and tested DrylANNd across 28 AmeriFlux eddy covariance sites (Fig. 1; Table S1), each overlapping the SMAP record (2015-present) by at least one year and consisting predominantly of natural vegetation. Based on the 1981-2010 TerraClimate annual precipitation (P) and potential evapotranspiration (PET) normals (Abatzoglou et al., 2018), most sites are semiarid ($0.2 \leq P/PET < 0.5$; *N*=16), with several others classified as arid ($0.03 \leq P/PET < 0.2$; *N*=6) and subhumid ($0.5 \leq P/PET \leq 0.75$; *N*=5). One site (Valles Caldera Mixed Conifer; US-Vcm) is slightly wetter than subhumid (P/PET = 0.85) due to its relatively high elevation (>3,000 m), but we include it here because it is part of the six-site New Mexico Elevation Gradient network (Anderson-Teixeira et al., 2011) and is frequently used in U.S. dryland flux research (e.g., Biederman et al.,

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

2017, 2016). The sites span a large latitudinal gradient (31.74°N to 46.69°N) and include six evergreen needleleaf forest (ENF), seven grassland (GRS), ten shrubland (SHB), and five savanna (SAV) sites.

We used a spike detection method to filter out sudden but temporary changes in half-hourly NEE, which can arise either from biophysical effects (e.g., sudden changes in turbulence) or from instrument error (Papale et al., 2006). Using REddyProc (Wutzler et al., 2018, 2020) in the R statistical computing environment (R Core Team, 2021), we then excluded half-hourly NEE observations that occurred during periods of low turbulence based on a seasonal friction velocity (U*) filter, defined empirically based on the U* distribution within each site and season using the moving point method (Wutzler et al., 2018; Papale et al., 2006), and gap-filled the missing data using a look-up table based on air temperature, shortwave radiation, and vapor pressure deficit (Papale et al., 2006). We also gap-filled latent heat flux estimates using the same method and converted latent heat flux to ET (in mm). We partitioned the half-hourly NEE into its component parts (GPP and ecosystem respiration) using the nighttime partitioning method (Papale et al., 2006; Reichstein et al., 2005), summed the half-hourly GPP, NEE, and ET to total daily fluxes, and calculated mean daily fluxes for each calendar month.

2.2 Remote sensing data

For each site, we obtained daily 500-meter multispectral surface reflectance from the MODIS Nadir Bidirectional Reflectance Distribution Function (BRDF)-Adjusted surface Reflectance (NBAR) product (MCD43A4) (Schaaf et al., 2002), excluding observations for which only fill values were provided (i.e., where a BRDF inversion could not be achieved) or where reflectance was below zero. For regional-scale prediction, we used the 0.05° MODIS Climate Modeling Grid (CMG) version of the NBAR product (MCD43C4). From the seven-band surface reflectance, we calculated seven daily VIs (Table 1): the normalized difference vegetation index (NDVI) (Rouse et al., 1974; Tucker, 1979), enhanced vegetation index (EVI) (Huete et al., 2002), near-infrared reflectance of vegetation (NIR_v) (Badgley et al., 2017), kernel NDVI (kNDVI) (Camps-Valls et al., 2021), and three versions of the land surface water index (LSWI) (Gao, 1996; Xiao et al., 2004), each based on a different shortwave infrared band (centered at 1240, 1640, and 2130 nm, respectively). NDVI is the most widely used vegetation index, but it sometimes fails to capture temporal dynamics of carbon and water fluxes in drylands (Yan et al., 2019; Smith et al., 2019; Wang et al., 2022). EVI, which includes soil and atmospheric adjustment factors, is more robust to soil background reflectance (Huete et al., 1994, 2002). The relatively new NIR_v and kNDVI have not yet been widely tested, but NIR_v has shown strong performance for predicting seasonal variability of dryland GPP (Wang et al., 2022). LSWI is related to water content of the land surface due to the strong absorption of shortwave-infrared by water (Ceccato et al., 2001; Gao, 1996; Xiao et al., 2004).

We also obtained 1-km resolution LST estimates for each site from thermal infrared observations onboard both MODIS Terra (MOD11A1) and Aqua (MYD11A1) (Wan, 2014), derived via a view-angle-dependent split window algorithm (Wan and Dozier, 1996). Each product provides one nighttime and one daytime observation per day (Table 1), with local overpass times at approximately 10:30 am and 10:30 pm (Terra) and 2:30 am and 2:30 pm (Aqua). For regional prediction, we used the corresponding 0.05° resolution CMG product (MOD11C1/MYD11C1). Surface temperature is an important determinant of carbon and water fluxes because, among other reasons, both photosynthesis and respiration involve temperature-dependent enzymatic reactions (Farquhar et al., 1980; Atkin and Tjoelker, 2003) and because it is a key indicator of latent heat flux, which cools leaves and land surfaces (Bateni and Entekhabi, 2012). Importantly for dryland ecosystems, accuracy of the MOD11/MYD11 collection 6 LST retrievals over bare soil is considerably improved over previous versions due to inclusion of separate daytime and nighttime coefficients and an emissivity adjustment model (Wan, 2014).

To capture soil moisture and soil temperature, we used daily (0:00 UTC retrieval) 9-km resolution surface (0-5 cm depth) and rootzone (0-100 cm) soil moisture and soil temperature estimates from the SMAP Level 4 Soil Moisture (L4SM) “analysis update” product (Reichle et al., 2019) (Table 1). SMAP L4SM assimilates satellite-observed L-band (1.41 GHz) microwave brightness temperature (sensitive to moisture in the upper layers of the soil and vegetation) into a hydrological model forced with instrumental precipitation observations (Reichle et al., 2019). For site-level calibration and evaluation, we retrieved soil moisture and temperature from the grid cell nearest to each tower. For regional analyses, we resampled the SMAP L4SM products to the 0.05° MODIS CMG resolution using the nearest neighbor.

Since the relationships between vegetation indices and ecosystem function can vary among different vegetation types (Wang et al., 2022), we used the 2020 fractional cover of annual and perennial grasses and forbs, trees, shrubs, litter, and bare ground from Rangeland Analysis version 3 (Jones et al., 2018; Allred et al., 2021) as static predictors. Rangeland Analysis fractional cover fuses Landsat and MODIS surface reflectance at ~30-meter resolution. For site-level model development, we averaged the fractional covers for all pixels within a 500-meter buffer around each eddy covariance tower. For regional analysis, we reprojected and resampled the 30-meter fractional cover to 0.0005° (~50 meter) resolution and then averaged the fractional cover of the 10,000 pixels falling within each 0.05° pixel of the MODIS CMG grid.

From the daily MODIS and SMAP observations, we developed monthly composites of each variable. For the optical vegetation “greenness” indices (NDVI, kNDVI, EVI, and NIR_v), we used maximum value compositing of the valid daily observations within each month, consistent with theoretical and observational evidence that sources of noise in remotely sensed imagery (e.g., clouds, snow, and atmospheric effects) tend to reduce, not increase, the apparent greenness of the land surface (Viovy et al., 1992; Huete et al., 2002). For LST, LSWI, and SMAP, we averaged the daily observations within each month. In all cases, we only used monthly composites where at least 25% of daily observations were valid within the composite window (i.e., at least eight days of valid observations within a given month).

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

2.3 Model framework and initialization

We used feed-forward artificial neural networks (ANNs; Fig. 2) to jointly predict monthly GPP, NEE, and ET. ANNs are effective at finding underlying relationships within multidimensional and multisource datasets. They are particularly useful for estimating biophysical parameters because they support non-linearity, adaptivity to changes in the environment, and decision confidence (Mas and Flores, 2008; Jensen et al., 2009). Synthetic "neurons," where each neuron is a mathematical function, connect the neural network's input and output layers, often through "hidden" layers of intermediary functions. Importantly, ANNs are appropriate for multi-output regression problems, where a single model simultaneously produces predictions of multiple variables. Because the multi-neuron output layer of the neural network allows joint prediction of response variables, the ANN framework therefore implicitly preserves some biophysical connections between GPP and NEE, where GPP is the carbon input into the ecosystem, and between GPP and ET, which are coupled via plant stomata.

The DrylANNd model consists of an ensemble of ANNs, each with one input layer of 20 "neurons" (i.e., seven optical VIs, LST observations from four different times per day, three SMAP variables, and six static fractional cover classes), two hidden layers, and one output layer with three neurons (GPP, NEE, and ET). The sizes of the two hidden layers (L_1 and L_2) were determined based on the number of neurons in the input (N) and output (m) layers (Huang, 2003):

$$L_1 = \sqrt{(m+2)N} + 2\sqrt{N/(m+2)} = 14$$

$$L_2 = m\sqrt{N/(m+2)} = 6$$

Each ANN in the ensemble (§2d below) was initiated with randomly assigned weights and biases based on the Nguyen-Widrow method (Nguyen and Widrow, 1990) and with different random subsets of observations for model training (75%) and validation (25%). We trained the ANNs using the Levenberg-Marquardt algorithm, which performed faster than and at least as well as other training algorithms in early tests.

2.4 Model calibration, evaluation, and prediction

Our DrylANNd model consists of an ensemble of 560 individual ANNs, where each ANN in the ensemble was trained with a different combination of sites (always withholding data from one site to use for independent evaluation) and initialized with different weights and biases that connect one layer to the next. Specifically, we withheld each AmeriFlux site from model development for 20 of the ensemble members (28 sites × 20 models per site = 560 total models), training the ANNs for those ensemble members using the remaining 27 sites. We then predicted GPP, NEE, and ET for the withheld site using each of the 20 ensemble members, with the ensemble mean used as the best estimate of GPP, NEE, and ET. This ensures that model skill is assessed based on observations from a site that was new to the model, thus providing a representative measure of the model's ability to extrapolate to locations on which it was not trained. Using the full ensemble of 560 ANNs, we predicted monthly GPP, NEE, and ET using the 0.05° resolution remotely sensed data for all Western U.S. drylands from April 2015 (the first full month of SMAP data) through December 2020. We used the ensemble mean at each pixel as an estimate of GPP, NEE, and ET, with the 10th-90th percentiles used as estimates of uncertainty.

We evaluated model skill based on the coefficient of determination (R^2) and mean absolute error (MAE) between holdout model predictions and observations at each site. In addition to evaluating skill at monthly resolution, we also assessed DrylANNd's ability to capture both the mean seasonality (i.e., the mean monthly fluxes during the 2015-2020 period) and the (inter)annual variability at the 16 AmeriFlux sites that were operational over the full SMAP period (2015-2020). We assessed the ability of the model to capture the mean seasonal cycle using both the R^2 and the standard deviation ratio (SDR), i.e., the ratio of the standard deviations of the modeled mean seasonal cycle and the observed mean seasonal cycle (Smerdon et al., 2011), where SDR<1 indicates that the model underestimated the seasonality and SDR>1 indicates that the model overestimated the seasonality. To assess the ability of the model to capture (inter)annual variability, we calculated annual fluxes of each year during the April-October warm season. We specifically examined annual warm season, rather than calendar year, fluxes because 1) SMAP began operation on March 31, 2015 so there is not a continuous record of January-March fluxes over the full 2015-2020 period and 2) many high elevation and/or high latitude sites have extensive missing data during the cold season due to snow cover. As a benchmark for model skill, we also compared DrylANNd predictions to the MODIS GPP (MOD17A2HGF; Running et al., 2004) and ET (MOD16A2GF; Mu et al., 2007, 2011) products.

The "black box" nature of many machine learning methods (including ANNs) typically makes it challenging to examine the effect of any given input variable on model predictions. Here, we examined the importance of the MODIS and SMAP predictor variables for model skill in two ways. First, using the same leave-one-site-out calibration and evaluation procedure described above, we ran models based on each of the three classes of variables (MODIS VIs, MODIS LST, and SMAP soil moisture/temperature) individually and in all possible combinations and compared the predictive ability (R^2) of each model to the combined model with all three classes of variables together. Second, we tested the leverage of each time-varying predictor variable by repeatedly (100 times) randomly permuting each variable (thus destroying its information content) and re-running model predictions. The synthetic, noise-only permutations of each variable were drawn from a normal distribution with the same mean and variance as observed for that variable at that site. We calculated both the mean percent increase in MAE and the change in R^2 when each variable was replaced with noise-only permutations, thus estimating its leverage on model skill.

3 Results

3.1 Overall performance

At the monthly scale, the DrylANNd model explained more than 70% of the variation in GPP (Fig. 3a) and ET (Fig. 3e) for sites withheld from model training, but only about 35% of the variation in NEE (Fig. 3c). For GPP, the model performed best at shrubland sites (R^2~0.8), followed by savanna and grassland sites (R^2~0.7), and evergreen needleleaf sites (R^2~0.6) (Fig.

3b). However, while the model performed worst on average at evergreen needleleaf sites, where optical VIs struggle to capture seasonal dynamics in GPP (Wang et al., 2022), these sites also saw the greatest *improvement* over MODIS GPP estimates (R^2~0.6 vs. R^2~0.25) (Fig. 3b). DrylANNd outperformed MODIS GPP estimates at 19 sites, while MODIS outperformed DrylANNd at 7 sites (with nearly identical performance at two sites). Overall, however, DrylANNd showed improvement over MODIS for all four vegetation types (Fig. 3b). Model skill was considerably worse for NEE (Fig. 3c-d), with best performance at shrubland sites (R^2~0.7), followed by grasslands (R^2~0.4), and then by savannas (R^2~0.3) and evergreen needleleaf forests (R^2~0.1) (Fig. 3d). The DrylANNd model particularly excelled at capturing variation in ET (Fig. 3e-f), especially in shrublands ($R^2>0.8$), with $R^2>0.6$ on average across all vegetation types (Fig. 3f). Site-level R^2 varied from ~0.4-0.9, which represents an improvement (in many cases substantial) over MODIS ET estimates at 23 out of 26 sites and for all four vegetation types (Fig. 3f).

DrylANNd also effectively captured spatial variation of warm-season carbon and water fluxes across western U.S. drylands (Fig. 4). The model simulates realistic spatial gradients of GPP (Fig. 4a), NEE (Fig. 4c), and ET (Fig. 4e), with highest productivity and ET in the subhumid east and in high elevation “sky islands,” where cooler temperatures and more abundant precipitation provide a more favorable environment than the surrounding desert lowlands. Across the 16 AmeriFlux sites that completely overlap the SMAP observational period, DrylANNd captured 75-80% of the spatial variation in warm-season GPP and ET (Figs. 4b and 4f, respectively) with minimal bias (i.e., with predictions all falling along the 1:1 line). For NEE, the model captured ~50% of the spatial variation but with a negative bias indicating an overestimation of the carbon sink across the region, particularly for evergreen needleleaf forest sites (Fig. 4d).

3.2 Seasonality

At the 16 AmeriFlux sites that cover the full SMAP period (2015-2020), DrylANNd effectively captured the mean seasonality (i.e., mean monthly fluxes) of both GPP and ET across most sites, with $R^2 \geq 0.8$ at all but two sites for both GPP (Fig. 5) and ET (Fig. 6). This represents an improvement over MODIS GPP seasonality at 12/16 sites and MODIS ET at 13/16 sites. Likewise, DrylANNd effectively represented the seasonality of GPP and ET, with SDR closer to 1 at 10/16 sites for GPP (Fig. 5) and 15/16 sites for ET (Fig. 6). However, even though DrylANNd’s ET estimates come considerably closer to representing the seasonality of ET, they still underestimate the magnitude of seasonality of ET (SDR<1) at all but two sites (US-Mpj and US-Ses; Fig. 6), suggesting that the disproportionately-important “hot moments” of dryland water fluxes that occur during intermittent pulses of rainfall are still not completely captured by the model. Likewise, DrylANNd often misses the bimodal spring/summer growing seasons at many of the North American monsoon-influenced sites (Figs. 5-6, S1-S2). For example, DrylANNd correctly estimated the magnitude of monsoon-driven summer GPP and ET but underestimated the spring GPP and ET at the Santa Rita Grassland (US-SRG) site in southeastern Arizona (Fig. 5-6, S1-S2), and it missed the mid-summer suppression of GPP at the two Valles Caldera sites (US-Vcp and US-Vcm) in New Mexico (Fig. 5, S1).

While the seasonality of NEE was mostly well captured by DrylANNd ($R^2 \geq 0.6$ at most sites; Fig. S3), it tended to systematically underestimate NEE (i.e., overestimate the magnitude of the carbon sink) at several sites (Fig. S4), particularly the three forest sites (US-Me6, US-Vcm, and US-Vcp), with a smaller number of sites where NEE was slightly overestimated. In the most extreme case (US-Vcp), the mean bias in NEE exceeded –1 g C m^{-2} day^{-1}, with biases approaching or exceeding –2 g C m^{-2} day^{-1} during the June-September period when measured NEE at the site was near zero.

3.3 (Inter)annual variability

DrylANNd captured roughly 70% of the variability in annual warm-season GPP (Fig. 7a) and 66% of the variability in warm-season ET (Fig. 7c) with MAE of ~100 g C m^{-2} for GPP and ~50 mm for ET, a considerable improvement over the MODIS GPP and ET estimates. However, like many remote sensing estimates of GPP and ET (Smith et al., 2019; Biederman et al., 2017; Stocker et al., 2019), DrylANNd struggled to capture the interannual *variability* (i.e., deviations from site mean) of carbon and water fluxes. DrylANNd only captured 31% of the interannual variability of GPP (Fig. 7b), similar to that of MODIS GPP (R^2=0.33) though with a slope slightly (but not significantly) closer to 1. Like MODIS, DrylANNd also underestimated the magnitude of interannual variability of ET, with a slightly lower R^2 than MODIS but a slope slightly closer to 1 (Fig. 7d).

3.4 Variable importance

For all three response variables (GPP, NEE, and ET), models that included all three subsets of predictor variables (optical VIs, LST, and SMAP) performed best overall (Fig. 8a-c), though a combination of optical vegetation indices with SMAP soil moisture/temperature performed nearly as well for both GPP (Fig. 8a) and ET (Fig. 8c). Models based on VIs and/or LST performed worse than SMAP-based models but still achieved overall R^2 of ~0.4 for GPP and ~0.5 for ET (compared to R^2~0.7 for the models that included all predictor variables). The same was generally true for the models' ability to capture within-site temporal variability (i.e., anomalies relative to monthly site-means; Fig. 8d-f). For both GPP (Fig. 8d) and ET (Fig. 8f), the model based on all predictor variables performed best for capturing temporal variability, followed closely by the VI+SMAP model.

Compared to models based solely on optical VIs, the addition of SMAP soil moisture/temperature generally made the largest difference for model performance in grasslands and shrublands, while including LST estimates from MODIS thermal infrared made the largest difference for model performance in evergreen needleleaf forests (Fig. 8). In shrublands, the R^2 of VI+SMAP GPP (Fig. 8a) and ET (Fig. 8c) increased by more than 0.1 compared to the VI-only models, and it increased by more than 0.2 for NEE (Fig. 8b). In grasslands, the difference between the VI-only and VI+SMAP models was even larger, with R^2 increasing by more than 0.2 for all three flux variables (Fig. 8a-c). Generally, the addition of thermal data, however, offered little (if any) gains in model performance over either the VI-only or VI+SMAP models for either grassland or shrubland sites (Fig. 8a-c), though it did improve representation of flux *anomalies* at grassland sites compared to VI-only models (Fig. 8d-f). Conversely, across evergreen needleleaf forest sites, the VI+SMAP models performed slightly worse than the VI-only models for predicting monthly fluxes (Fig. 8a-c), suggesting that models trained on other evergreen needleleaf sites were overfit to the SMAP data and were less skillful at extrapolating to an "unseen" site than a model based solely on optical VIs. However, including SMAP soil moisture/temperature did slightly improve the ability of the models to predict monthly flux *anomalies* at evergreen needleleaf forests (Fig. 8d-f). Including thermal data, on the other hand, improved predictions of both monthly fluxes and monthly flux anomalies in evergreen needleleaf forests compared to VI-only models. For savanna sites, both LST and SMAP soil moisture/temperature improved predictions of monthly fluxes compared to VI-only models (Fig. 8a-c), though SMAP was far more important for predicting flux anomalies (Fig. 8d-f).

Interestingly, while the LST-only models performed the worst overall (Fig. 8), the models based on all variables assigned high leverage to LST for all three response variables, wherein a random permutation of LST led to large increases in MAE and decreases in R^2 (Fig. 9). For GPP, the four diurnal LST observations (particularly daytime Terra LST), along with NDVI and NIR_v, tended to have the highest "leverage" on model skill (Fig. 9a-b). There was little variation in leverage among the predictor variables for NEE (Fig. 9c-d), with random permutations generally leading to relatively low and site-specific changes in both MAE and R^2. However, nighttime LST observations had the greatest leverage over the MAE of NEE predictions (Fig. 9c), while NDVI/NIR_v and all four daily LST observations had the most leverage over predictive R^2 (Fig. 9d). For ET, a combination of all three classes of variables (vegetation indices, LST, and soil moisture/temperature) contributed positively to model skill, with NDVI/NIR_v, daytime and nighttime Terra LST, and SMAP surface soil moisture generally holding the highest leverage over model skill (Fig. 9e-f).

4 Discussion

4.1 Model skill, strengths, and shortcomings

Here, we developed and evaluated a data-driven, machine learning-based approach for modeling monthly carbon (GPP, NEE) and water (ET) fluxes in U.S. drylands using multi-source satellite remote sensing. Our DrylANNd model incorporated information from the optical, thermal, and microwave domains, including newer optical VIs that have shown promise in drylands (i.e. NIRv; Wang et al., 2022), daily land surface temperature observations from multiple times per day, and estimates of surface/rootzone soil moisture and soil temperature. DrylANNd performed particularly well at monthly and seasonal (i.e., mean monthly) time scales, representing a considerable improvement over MODIS GPP and ET estimates across most eddy covariance sites and all vegetation types (Fig. 3).

DrylANNd particularly excelled at capturing monthly (Fig. 3e-f), seasonal (Fig. 6), and spatial (Fig. 4e-f) variation in ET. Given the importance of ET for linking the carbon, water, and energy cycles (Fisher et al., 2017), accurate ET estimates are critical for understanding and monitoring global ecosystem function, especially in drylands where remote sensing of ET is particularly challenging (Smith et al., 2019; Fisher et al., 2017). By contrast, NEE proved more challenging to estimate than either GPP or ET (Fig. 3c-d, 4c-d), likely because many processes involved in ecosystem respiration cannot be easily represented with satellite data. While heterotrophic and autotrophic respiration rates are strongly dependent on temperature (Atkin and Tjoelker, 2003) and soil moisture (Moyano et al., 2013), which can be captured by MODIS LST and SMAP soil moisture/temperature, they also depend on microbial community composition, substrate availability, and root biomass that are not visible to satellite sensors. DrylANNd performed moderately well at capturing the seasonality (Fig. S3) and spatial variation (Fig. 4c-d) in NEE but tended to systematically overestimate the magnitude of net carbon uptake in U.S. drylands (Fig. 4d, S3), particularly in evergreen needleleaf forests. Many dryland sites have a net carbon balance near zero and can flip between being sources of and sinks for CO_2 in any given year (Biederman et al., 2018; Scott et al., 2009, 2010, 2015), so even small errors or biases in NEE estimates can have large effects on dryland carbon budgeting and monitoring.

While DrylANNd captured monthly, seasonal (i.e., mean monthly), and spatial variation of GPP and ET with fidelity, it struggled to predict interannual variability (Fig 7). This is a common issue for satellite-based models applied in dryland ecosystems (Biederman et al., 2017; Smith et al., 2019; Stocker et al., 2019; Barnes et al., 2021), partly due to the prevalence of "hot moments" (i.e., short periods of high biogeochemical activity) that are disproportionately important to time-averaged carbon and water fluxes in drylands (Kannenberg et al., 2020). While DrylANNd has relatively little systematic bias at capturing low extremes in monthly GPP (Fig. 3a) and ET (Fig. 3e), it tended to underestimate the high extremes. DrylANNd's monthly resolution may smooth the intensity of these short but impactful "hot moments," leading to systematic underestimation of monthly high extremes which also propagates to longer timescales, with DrylANNd clearly underestimating the high extremes in interannual variability of warm season GPP (Fig. 7a) and ET (Fig. 7c). Improving estimates of interannual variability of dryland systems may therefore require models that operate at finer temporal resolutions (e.g., daily) to adequately represent short, intense periods of pulse-driven dryland vegetation activity.

Despite the challenges in capturing interannual variability, the ANN machine learning approach used here has several key benefits. First, because it is a data-driven model based solely on remote sensing products with short latencies, DrylANNd would be relatively easy to operationalize at a large scale and in near real time. Second, the ensemble approach allows for intuitive estimates of uncertainty, which are critical for many applications (e.g., ecological forecasting) but which are rarely provided (Dietze et al., 2018). Finally, neural networks allow joint modeling of multiple response variables, providing the means both to efficiently generate multiple indicators of ecosystem activity and to partially preserve the physical connections between GPP and NEE and between GPP and ET, which is relatively rare for remote sensing-based models. The MODIS and SMAP carbon products, for example, provide joint estimates of GPP and net primary production (Running et al., 2004) and GPP and NEE (Jones et al., 2017), respectively, but neither provides estimates of ET that are coupled to GPP. Zhang et al. (2016), on the other hand, provide coupled estimates of GPP and ET using static, biome-specific water-use efficiencies, but this approach does not provide estimates of downstream plant or ecosystem carbon balances, nor does the model allow for the dynamic changes in water-use efficiency that can occur in response to pulses of rainfall or variation in vapor pressure deficit (Roby et al., 2020).

4.2 Benefits of multi-source remote sensing

Previous work has highlighted the potential for combining multiple remote sensing proxies to improve the representation of vegetation dynamics (Stavros et al., 2017; Smith et al., 2019), and our results support this conclusion. All three classes of remote sensing variables (optical, thermal, and microwave) contributed positively to model skill (Figs. 8-9). In particular, the inclusion of SMAP soil moisture/temperature resulted in large gains in model skill (Fig. 8), with the VI+SMAP models performing substantially better than the VI-only models, especially for GPP (Fig. 8a), for which the monthly R^2 improved from ~0.4 to nearly 0.7 when adding SMAP variables as predictors. This is consistent with site-level research showing the importance of soil moisture for dryland carbon and water fluxes (Novick et al., 2016; Stocker et al., 2018). The inclusion of soil moisture and soil temperature was particularly important for predicting carbon and water fluxes in grassland, shrubland, and savanna systems.

While the LST-only models usually performed worst of all model subsets, and the monthly VI+LST models barely outperformed the VI-only models (Fig. 8a-c), the inclusion of LST was particularly useful for improving predictions of monthly *anomalies* (i.e., deviations from the mean monthly flux of a given site) of all three flux variables, with R^2 increasing by ~0.1 when adding LST to the VI-only models (Fig. 8d-f). When LST was included in the models, it also had very high leverage over model skill (Fig. 9), indicating that it provided unique information not captured by other remote sensing sources. Among the different vegetation types, thermal data was particularly important for evergreen needleleaf forests (Fig. 8), where optical VIs struggle to capture both spatial and temporal dynamics of GPP (Wang et al., 2022) and where growing seasons are more likely to be limited by low temperatures than in warmer and drier grasslands, shrublands, and savannas. Daytime LST was particularly important for predicting GPP and ET, possibly indicating that it is an effective indicator of moisture stress (Javadian et al., 2022; Still et al., 2021), while nighttime LST was particularly important for predicting NEE, consistent with previous work showing the importance of nighttime temperature for respiration and overall carbon balance (Anderegg et al., 2015). LST also has the benefit of a much longer operational record and much finer spatial resolution than SMAP, so models based on a combination of VIs and LST could allow longer records of dryland responses to hydroclimatic variability and change.

Somewhat surprisingly, NDVI held higher leverage over model predictions than most other predictor variables (Fig. 9), despite previous research documenting significant flaws in its ability to track dryland GPP (Wang et al., 2022; Yan et al., 2019). The relatively new NIR_v index also held very high leverage, consistent with recent work showing that it is particularly effective at tracking GPP variation in low productivity, sparsely-vegetated grasslands and shrublands (Wang et al., 2022). EVI, on the other hand, was assigned very low leverage across nearly all sites and all fluxes, despite generally being regarded as an

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

improvement over NDVI in drylands due to the inclusion of a soil background adjustment factor (Smith et al., 2019). This could plausibly result from EVI containing similar information as other optical vegetation indices and thus being assigned less weight in the models. Wang et al. (2022), for example, found that the soil-adjusted vegetation index (SAVI), on which EVI was partly based (Huete et al., 1994, 1997, 2002), performed very similarly to but slightly worse than NIR_v for tracking seasonality and spatial variability of dryland GPP.

4.3 Priorities for future dryland model development

The DrylANNd model's limitations could be ameliorated by incorporating additional remote sensing data that capture other aspects of dryland ecosystem function. Solar-induced fluorescence, for example, effectively tracks vegetation activity in drylands (Smith et al., 2018; Wang et al., 2022), particularly in dry evergreen needleleaf forests where reflectance-based optical VIs tend to perform poorly (Magney et al., 2019; Wang et al., 2022). However, satellite-based SIF estimates suffer from coarse spatial and temporal resolutions (e.g., GOME-2; Joiner et al., 2013), discontinuous spatial coverage (e.g., OCO-2 and OCO-3; Sun et al., 2018), or an even shorter period of record than SMAP (e.g., TROPOMI; Köhler et al., 2018). Fusions of satellite SIF data with MODIS surface reflectance (e.g., Zhang et al., 2018) overcome some of these limitations but would likely inherit many of the same flaws as reflectance-based optical VIs since they are based on the same surface reflectance data. As SIF temporal and spatial resolution improves, it will likely become increasingly useful for dryland carbon and water modeling. Gravimetric estimates of total terrestrial water storage (Andersen et al., 2005; Humphrey et al., 2018) could also improve the representation of deeper moisture, which can be an essential water source for deep-rooted trees in semiarid systems (Rempe and Dietrich, 2018; McCormick et al., 2021). However, like the longer-term SIF measurements, estimates of total water storage are limited to very coarse (0.5°) spatial resolution and monthly frequency, which would preclude use in higher frequency estimates of carbon and water fluxes.

5 Conclusions

Here, we developed and evaluated a machine learning approach (DrylANNd) for jointly modeling of key carbon and water fluxes (GPP, NEE, and ET) specifically for drylands of the western U.S. using a combination of satellite optical vegetation indices, multitemporal thermal infrared, and passive microwave-based soil moisture and soil temperature. Longstanding challenges in current multispectral satellite-based estimation of dryland carbon and water fluxes are the result of several interacting issues, including poor representation of soil moisture stress, decoupling between "greenness" and plant physiology, high soil background reflectance in open canopies, and limited representation of dryland calibration and validation sites available for model training and testing. Our approach partially addresses these limitations of previous satellite carbon and water flux estimates in drylands:

- Soil moisture is explicitly included in the model rather than relying on the covariance between VPD and soil moisture or water-sensitive vegetation indices as proxies of moisture stress.

- The model includes new vegetation indices (e.g., NIR_v) that show promise in capturing dryland seasonality (Wang et al., 2022), along with satellite thermal and microwave observations representing temperature and moisture stress impacts on ecosystem fluxes.
- The model is trained specifically for dryland ecosystems based on an extensive network of 28 eddy covariance sites spanning a large latitudinal and arid-to-subhumid gradient in the western U.S.

We found that this approach effectively captures monthly, seasonal, and spatial variation in GPP and, especially, ET through both space and time, though it still underestimates the magnitude of interannual variability of carbon and water fluxes. DrylANNd was less effective at capturing NEE than GPP or ET, likely because respiratory processes are largely invisible to satellite sensors, with the magnitude of dryland carbon sinks overestimated particularly at evergreen needleleaf sites. Compared to models based solely on optical vegetation indices, the inclusion of SMAP soil moisture/temperature was crucial for improving estimates of both the magnitudes and temporal variabilities of all three fluxes, especially in dry grasslands and shrublands of the western U.S. On the other hand, the addition of multitemporal thermal observations improved flux estimates in evergreen needleleaf forests, where optical vegetation indices have traditionally struggled to capture GPP dynamics. Drylands play important roles both in the global carbon cycle (Poulter et al., 2014; Ahlström et al., 2015) and in ecosystem services supporting a large human population (Reynolds et al., 2007; Bestelmeyer et al., 2015), and DrylANNd significantly improves our ability to quantify carbon and water fluxes in these ecosystems.

Data and code availability. Code for all modeling and analysis is available at https://github.com/mpdannenberg/drylANNd, and 0.05° monthly DrylANNd GPP, NEE, and ET estimates for the western U.S. are archived and publicly available as NetCDF files distributed under a creative commons license at Iowa Research Online (https://doi.org/10.25820/data.006185).

Author contributions. Conceptualization: MPD, MLB, WKS, RLS, and JAB. Formal analysis: MPD and XW. Funding acquisition: MPD, MLB, WKS, and XW. Methodology: MPD, MLB, MRJ, and SKM. Software: MPD. Visualization: MPD. Writing – original draft preparation: MPD, MLB, and MRJ. Writing – review & editing: all authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. MPD, MLB, WKS, and MRJ were supported by the NASA SMAP Science Team (grant number 80NSSC20K1805), and XW was supported by the NASA FINESST program (grant number 80NSSC19K1335). We thank AmeriFlux and the tower PIs for making eddy covariance data publicly available. Any use of firm, product or trade names is for descriptive purposes only and does not imply endorsement by the U.S. Government. USDA is an equal-opportunity employer and provider.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015, Sci. Data, 5, 1–12, https://doi.org/10.1038/sdata.2017.191, 2018.

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., and Zeng, N.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink, Science, 348, 895–899, https://doi.org/10.1002/2015JA021022, 2015.

Allred, B. W., Bestelmeyer, B. T., Boyd, C. S., Brown, C., Davies, K. W., Duniway, M. C., Ellsworth, L. M., Erickson, T. A., Fuhlendorf, S. D., Griffiths, T. V., Jansen, V., Jones, M. O., Karl, J., Knight, A., Maestas, J. D., Maynard, J. J., McCord, S. E., Naugle, D. E., Starns, H. D., Twidwell, D., and Uden, D. R.: Improving Landsat predictions of rangeland fractional cover with multitask learning and uncertainty, Methods Ecol. Evol., 12, 841–849, https://doi.org/10.1111/2041-210X.13564, 2021.

Anderegg, W. R. L., Ballantyne, A. P., Smith, W. K., Majkut, J., Rabin, S., and Beaulieu, C.: Tropical nighttime warming as a dominant driver of variability in the terrestrial carbon sink, Proc. Natl. Acad. Sci., 112, 15591–15596, https://doi.org/10.1073/pnas.1521479112, 2015.

Andersen, O. B., Seneviratne, S. I., Hinderer, J., and Viterbo, P.: GRACE-derived terrestrial water storage depletion associated with the 2003 European heat wave, Geophys. Res. Lett., 32, https://doi.org/10.1029/2005GL023574, 2005.

Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., and Litvak, M. E.: Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico, Glob. Chang. Biol., 17, 410–424, https://doi.org/10.1111/j.1365-2486.2010.02269.x, 2011.

Anderson, M. C., Allen, R. G., Morse, A., and Kustas, W. P.: Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources, Remote Sens. Environ., 122, 50–65, https://doi.org/10.1016/j.rse.2011.08.025, 2012.

Atkin, O. K. and Tjoelker, M. G.: Thermal acclimation and the dynamic response of plant respiration to temperature, Trends Plant Sci., 8, 343–351, https://doi.org/10.1016/S1360-1385(03)00136-5, 2003.

Ault, T. R.: On the essentials of drought in a changing climate, Science, 368, 256–260, 2020.

Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, Sci. Adv., 3, e1602244, https://doi.org/10.1126/sciadv.1602244, 2017.

Barnes, M. L., Farella, M. M., Scott, R. L., Moore, D. J. P., Ponce-Campos, G. E., Biederman, J. A., MacBean, N., Litvak, M. E., and Breshears, D. D.: Improved dryland carbon flux predictions with explicit consideration of water-carbon coupling, Commun. Earth Environ., 2, 1–9, https://doi.org/10.1038/s43247-021-00308-2, 2021.

Bateni, S. M. and Entekhabi, D.: Relative efficiency of land surface energy balance components, Water Resour. Res., 48, 1–8, https://doi.org/10.1029/2011WR011357, 2012.

Bestelmeyer, B. T., Okin, G. S., Duniway, M. C., Archer, S. R., Sayre, N. F., Williamson, J. C., and Herrick, J. E.: Desertification, land use, and the transformation of global drylands, Front. Ecol. Environ., 13, 28–36, https://doi.org/10.1890/140162, 2015.

Biederman, J. A., Scott, R. L., Goulden, M. L., Vargas, R., Litvak, M. E., Kolb, T. E., Yepez, E. A., Oechel, W. C., Blanken, P. D., Bell, T. W., Garatuza-Payan, J., Maurer, G. E., Dore, S., and Burns, S. P.: Terrestrial carbon balance in a drier world: The effects of water availability in southwestern North America, Glob. Chang. Biol., 22, 1867–1879, https://doi.org/10.1111/gcb.13222, 2016.

Biederman, J. A., Scott, R. L., Bell, T. W., Bowling, D. R., Dore, S., Garatuza-Payan, J., Kolb, T. E., Krishnan, P., Krofcheck, D. J., Litvak, M. E., Maurer, G. E., Meyers, T. P., Oechel, W. C., Papuga, S. A., Ponce-Campos, G. E., Rodriguez, J. C., Smith, W. K., Vargas, R., Watts, C. J., Yepez, E. A., and Goulden, M. L.: CO2 exchange and evapotranspiration across dryland ecosystems of southwestern North America, Glob. Chang. Biol., 23, 4204–4221, https://doi.org/10.1111/gcb.13686, 2017.

Biederman, J. A., Scott, R. L., Arnone, J. A., Jasoni, R. L., Litvak, M. E., Moreo, M. T., Papuga, S. A., Ponce-Campos, G. E., Schreiner-McGraw, A. P., and Vivoni, E. R.: Shrubland carbon sink depends upon winter water availability in the warm deserts of North America, Agric. For. Meteorol., 249, 407–419, https://doi.org/10.1016/j.agrformet.2017.11.005, 2018.

Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., Mahecha, M. D., Muñoz-Marí, J., García-Haro, F. J., Guanter, L., Jung, M., Gamon, J. A., Reichstein, M., and Running, S. W.: A unified vegetation index for quantifying the terrestrial biosphere, Sci. Adv., 7, 1–11, https://doi.org/10.1126/sciadv.abc7447, 2021.

Cayan, D. R., Das, T., Pierce, D. W., Barnett, T. P., Tyree, M., and Gershunov, A.: Future dryness in the southwest US and the hydrology of the early 21st century drought., Proc. Natl. Acad. Sci. U. S. A., 107, 21271–21276, https://doi.org/10.1073/pnas.0912391107, 2010.

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., and Grégoire, J.-M.: Detecting vegetation leaf water content using reflectance in the optical domain, Remote Sens. Environ., 77, 22–33, https://doi.org/10.1016/S0034-4257(01)00191-2, 2001.

Cook, B. I., Ault, T. R., and Smerdon, J. E.: Unprecedented 21st century drought risk in the American Southwest and Central Plains, Sci. Adv., 1, e1400082, https://doi.org/10.1126/sciadv.1400082, 2015.

Cook, B. I., Mankin, J. S., Marvel, K., Williams, A. P., Smerdon, J. E., and Anchukaitis, K. J.: Twenty-first century drought projections in the CMIP6 forcing scenarios, Earth's Futur., e2019EF001461, https://doi.org/10.1029/2019ef001461, 2020.

Dannenberg, M., Wang, X., Yan, D., and Smith, W.: Phenological characteristics of global ecosystems based on optical, fluorescence, and microwave remote sensing, Remote Sens., 12, https://doi.org/10.3390/rs12040671, 2020.

Dannenberg, M. P., Song, C., Hwang, T., and Wise, E. K.: Empirical evidence of El Niño-Southern Oscillation influence on land surface phenology and productivity in the western United States, Remote Sens. Environ., 159, 167–180, 2015.

Dannenberg, M. P., Smith, W. K., Zhang, Y., Song, C., Huntzinger, D. N., and Moore, D. J. P.: Large-scale reductions in terrestrial carbon uptake following central Pacific El Niño, Geophys. Res. Lett., 48, 1–11, https://doi.org/10.1029/2020GL092367, 2021.

Dannenberg, M. P., Yan, D., Barnes, M. L., Smith, W. K., Johnston, M. R., Scott, R. L., Biederman, J. A., Knowles, J. F., Wang, X., Duman, T., Litvak, M. E., Kimball, J. S., Williams, A. P., and Zhang, Y.: Exceptional heat and atmospheric dryness amplified losses of primary production during the 2020 U.S. Southwest hot drought, Glob. Chang. Biol., 28, 4794–4806, https://doi.org/10.1111/gcb.16214, 2022.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., and White, E. P.: Iterative near-term ecological forecasting: Needs, opportunities, and challenges, Proc. Natl. Acad. Sci. U. S. A., 115, 1424–1432, https://doi.org/10.1073/pnas.1710231115, 2018.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, 1980.

Fisher, J. B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M. F., Hook, S., Baldocchi, D., Townsend, P. A., Kilic, A., Tu, K., Miralles, D. D., Perret, J., Lagouarde, J.-P. P., Waliser, D., Purdy, A. J., French, A., Schimel, D., Famiglietti, J. S., Stephens, G., and Wood, E. F.: The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, Water Resour. Res., 53, 2618–2626, https://doi.org/10.1002/2016WR020175, 2017.

Gao, B.: NDWI--a normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sens. Environ., 58, 257–266, 1996.

Guan, K., Wu, J., Kimball, J. S., Anderson, M. C., Frolking, S., Li, B., Hain, C. R., and Lobell, D. B.: The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields, Remote Sens. Environ., 199, 333–349, https://doi.org/10.1016/j.rse.2017.06.043, 2017.

Haughton, N., Abramowitz, G., De Kauwe, M. G., and Pitman, A. J.: Does predictability of fluxes vary between FLUXNET sites?, Biogeosciences, 15, 4495–4513, https://doi.org/10.5194/bg-15-4495-2018, 2018.

Heinsch, F. A., Zhao, M., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., Bolstad, P. V., Cook, B. D., Desai, A. R., Ricciuto, D. M., Law, B. E., Oechel, W. C., Kwon, H., Luo, H., Wofsy, S. C., Dunn, A. L., Munger, J. W., Baldocchi, D. D., Xu, L., Hollinger, D. Y., Richardson, A. D., Stoy, P. C., Siqueira, M. B. S., Monson, R. K., Burns, S. P., and Flanagan, L. B.: Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations, IEEE Trans. Geosci. Remote Sens., 44, 1908–1925, https://doi.org/10.1109/TGRS.2005.853936, 2006.

Huang, G. Bin: Learning capability and storage capacity of two-hidden-layer feedforward networks, IEEE Trans. Neural Networks, 14, 274–281, https://doi.org/10.1109/TNN.2003.809401, 2003.

Huang, J., Yu, H., Guan, X., Wang, G., and Guo, R.: Accelerated dryland expansion under climate change, Nat. Clim. Chang., 6, 166–171, https://doi.org/10.1038/nclimate2837, 2016.

Huang, J., Yu, H., Dai, A., Wei, Y., and Kang, L.: Drylands face potential threat under 2°C global warming target, Nat. Clim. Chang., 7, 417–422, https://doi.org/10.1038/nclimate3275, 2017.

Huete, A. R.: A soil-adjusted vegetation index (SAVI), Remote Sens. Environ., 25, 295–309, https://doi.org/10.1016/0034-4257(88)90106-X, 1988.

Huete, A. R. and Jackson, R. D.: Suitability of spectral indices for evaluating vegetation characteristics on arid rangelands, Remote Sens. Environ., 23, 213–232, https://doi.org/10.1016/0034-4257(87)90038-1, 1987.

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

Huete, A. R., Justice, C., and Liu, H.: Development of vegetation and soil indices for MODIS-EOS, Remote Sens. Environ., 49, 224–234, https://doi.org/10.1016/0034-4257(94)90018-3, 1994.

Huete, A. R., Liu, H. Q., Batchily, K., and van Leeuwen, W.: A comparison of vegetation indices over a global set of TM images for EOS-MODIS, Remote Sens. Environ., 59, 440–451, 1997.

Huete, A. R., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ., 83, 195–213, https://doi.org/10.1016/S0034-4257(02)00096-2, 2002.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., and Seneviratne, S. I.: Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage, Nature, 560, 628–631, https://doi.org/10.1038/s41586-018-0424-4, 2018.

Huxman, T. E., Snyder, K. A., Tissue, D., Leffler, A. J., Ogle, K., Pockman, W. T., Sandquist, D. R., Potts, D. L., and Schwinning, S.: Precipitation pulses and carbon fluxes in semiarid and arid ecosystems, Oecologia, 141, 254–268, https://doi.org/10.1007/s00442-004-1682-4, 2004.

Javadian, M., Smith, W. K., Lee, K., Knowles, J. F., Scott, R. L., Fisher, J. B., Moore, D. J. P., van Leeuwen, W. J. D., Barron-Gafford, G., and Behrangi, A.: Canopy temperature Is regulated by ecosystem structural traits and captures the ecohydrologic dynamics of a semiarid mixed conifer forest site, J. Geophys. Res. Biogeosciences, 127, 1–15, https://doi.org/10.1029/2021JG006617, 2022.

Jensen, R. R., Hardin, P. J., and Yu, G.: Artificial neural networks and remote sensing, Geogr. Compass, 3, 630–646, https://doi.org/10.1111/j.1749-8198.2008.00215.x, 2009.

Joiner, J., Guanter, L., Lindstrot, R., Voigt, M., Vasilkov, A. P., Middleton, E. M., Huemmrich, K. F., Yoshida, Y., and Frankenberg, C.: Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, Atmos. Meas. Tech., 6, 2803–2823, https://doi.org/10.5194/amt-6-2803-2013, 2013.

Jones, L. A., Kimball, J. S., Reichle, R. H., Madani, N., Glassy, J., Ardizzone, J. V., Colliander, A., Cleverly, J., Desai, A. R., Eamus, D., Euskirchen, E. S., Hutley, L., Macfarlane, C., and Scott, R. L.: The SMAP Level 4 Carbon product for monitoring ecosystem land-atmosphere CO2 exchange, IEEE Trans. Geosci. Remote Sens., 55, 6517–6532, https://doi.org/10.1109/TGRS.2017.2729343, 2017.

Jones, M. O., Allred, B. W., Naugle, D. E., Maestas, J. D., Donnelly, P., Metz, L. J., Karl, J., Smith, R., Bestelmeyer, B., Boyd, C., Kerby, J. D., and McIver, J. D.: Innovation in rangeland monitoring: annual, 30 m, plant functional type percent cover maps for U.S. rangelands, 1984–2017, Ecosphere, 9, https://doi.org/10.1002/ecs2.2430, 2018.

Kannenberg, S. A., Bowling, D. R., and Anderegg, W. R. L.: Hot moments in ecosystem fluxes: High GPP anomalies exert outsized influence on the carbon cycle and are differentially driven by moisture availability across biomes, Environ. Res. Lett., 15, https://doi.org/10.1088/1748-9326/ab7b97, 2020.

Köhler, P., Frankenberg, C., Magney, T. S., Guanter, L., Joiner, J., and Landgraf, J.: Global retrievals of solar-induced chlorophyll fluorescence with TROPOMI: first results and intersensor comparison to OCO-2, Geophys. Res. Lett., 45, 10,456-10,463, https://doi.org/10.1029/2018GL079031, 2018.

MacBean, N., Scott, R. L., Biederman, J. A., Peylin, P., Kolb, T., Litvak, M. E., Krishnan, P., Meyers, T. P., Arora, V. K., Bastrikov, V., Goll, D., Lombardozzi, D. L., Nabel, J. E. M. S., Pongratz, J., Sitch, S., Walker, A. P., Zaehle, S., and Moore, D. J. P.: Dynamic global vegetation models underestimate net CO2 flux mean and inter-annual variability in dryland ecosystems, Environ. Res. Lett., 16, 094023, https://doi.org/10.1088/1748-9326/ac1a38, 2021.

Magney, T. S., Bowling, D. R., Logan, B. A., Grossmann, K., Stutz, J., Blanken, P. D., Burns, S. P., Cheng, R., Garcia, M. A., Köhler, P., Lopez, S., Parazoo, N. C., Raczka, B., Schimel, D., and Frankenberg, C.: Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence, Proc. Natl. Acad. Sci. U. S. A., 116, 11640–11645, https://doi.org/10.1073/pnas.1900278116, 2019.

Mas, J. F. and Flores, J. J.: The application of artificial neural networks to the analysis of remotely sensed data, Int. J. Remote Sens., 29, 617–663, https://doi.org/10.1080/01431160701352154, 2008.

McCormick, E. L., Dralle, D. N., Hahm, W. J., Tune, A. K., Schmidt, L. M., Chadwick, K. D., and Rempe, D. M.: Widespread woody plant use of water stored in bedrock, Nature, 597, 225–229, https://doi.org/10.1038/s41586-021-03761-3, 2021.

Moyano, F. E., Manzoni, S., and Chenu, C.: Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models, Soil Biol. Biochem., 59, 72–85, https://doi.org/10.1016/j.soilbio.2013.01.002, 2013.

Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sens. Environ., 111, 519–536, https://doi.org/10.1016/j.rse.2006.07.007, 2007.

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115, 1781–1800, https://doi.org/10.1016/j.rse.2011.02.019, 2011.

Nguyen, D. and Widrow, B.: Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights, in: 1990 IJCNN International Joint Conference on Neural Networks, 21–26, 1990.

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., and Phillips, R. P.: The increasing importance of atmospheric demand for ecosystem water and carbon fluxes, Nat. Clim. Chang., 6, 1023–1027, https://doi.org/10.1038/nclimate3114, 2016.

Olden, J. D., Lawler, J. J., and Poff, N. L.: Machine learning methods without tears: a primer for ecologists., Q. Rev. Biol., 83, 171–93, 2008.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: Algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, https://doi.org/10.5194/bg-3-571-2006, 2006.

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., and van der Werf, G. R.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, Nature, 509, 600–603, https://doi.org/10.1038/nature13376, 2014.

R Core Team: R: A Language and Environment for Statistical Computing, http://www.r-project.org, 2021.

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger, J., Starks, P., and Walker, J. P.: Version 4 of the SMAP Level-4 Soil Moisture algorithm and data product, J. Adv. Model. Earth Syst., 11, 3106–3130, https://doi.org/10.1029/2019MS001729, 2019.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm, Glob. Chang. Biol., 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Rempe, D. M. and Dietrich, W. E.: Direct observations of rock moisture, a hidden component of the hydrologic cycle, Proc. Natl. Acad. Sci. U. S. A., 115, 2664–2669, https://doi.org/10.1073/pnas.1800141115, 2018.

Reynolds, J. F., Smith, D. M. S., Lambin, E. F., Turner, B. L., Mortimore, M., Batterbury, S. P. J., Downing, T. E., Dowlatabadi, H., Fernández, R. J., Herrick, J. E., Huber-Sannwald, E., Jiang, H., Leemans, R., Lynam, T., Maestre, F. T., Ayarza, M., and Walker, B.: Global Desertification: Building a Science for Dryland Development, Science, 316, 847–851, https://doi.org/10.1126/science.1131634, 2007.

Roby, M. C., Scott, R. L., and Moore, D. J. P.: High vapor pressure deficit decreases the productivity and water use efficiency of rain-induced pulses in semiarid ecosystems, J. Geophys. Res. Biogeosciences, 125, e2020JG005665, https://doi.org/10.1029/2020JG005665, 2020.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C.: Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. NASA/GSFC Type III Final Report, Greenbelt, MD, 371 pp., 1974.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., and Hashimoto, H.: A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production, Bioscience, 54, 547–560, https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., D'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sens. Environ., 83, 135–148, https://doi.org/10.1016/S0034-4257(02)00091-3, 2002.

Scott, R. L., Jenerette, G. D., Potts, D. L., and Huxman, T. E.: Effects of seasonal drought on net carbon dioxide exchange from a woody-plant-encroached semiarid grassland, J. Geophys. Res., 114, G04004, https://doi.org/10.1029/2008JG000900, 2009.

Scott, R. L., Hamerlynck, E. P., Jenerette, G. D., Moran, M. S., and Barron-Gafford, G. A.: Carbon dioxide exchange in a semidesert grassland through drought-induced vegetation change, J. Geophys. Res., 115, G03026, https://doi.org/10.1029/2010JG001348, 2010.

Scott, R. L., Biederman, J. A., Hamerlynck, E. P., and Barron-Gafford, G. A.: The carbon balance pivot point of southwestern U.S. semiarid ecosystems: Insights from the 21st century drought, J. Geophys. Res. Biogeosciences, 120, 2612–2624, https://doi.org/10.1002/2015JG003181, 2015.

Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Bolstad, P. V., Flanagan, L. B., Goldstein, A. H., Hollinger, D. Y., Misson, L., Monson, R. K., Oechel, W. C., Schmid, H. P., Wofsy, S. C., and Xu, L.: A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS, Remote Sens. Environ., 112, 1633–1646, https://doi.org/10.1016/j.rse.2007.08.004, 2008.

Smerdon, J. E., Kaplan, A., Zorita, E., González-Rouco, J. F., and Evans, M. N.: Spatial performance of four climate field reconstruction methods targeting the Common Era, Geophys. Res. Lett., 38, L11705, https://doi.org/10.1029/2011GL047372, 2011.

Smith, W. K., Biederman, J. A., Scott, R. L., Moore, D. J. P., He, M., Kimball, J. S., Yan, D., Hudson, A., Barnes, M. L., MacBean, N., Fox, A. M., and Litvak, M. E.: Chlorophyll fluorescence better captures seasonal and interannual gross primary productivity dynamics across dryland ecosystems of southwestern North America, Geophys. Res. Lett., 45, 748–757, https://doi.org/10.1002/2017GL075922, 2018.

Smith, W. K., Dannenberg, M. P., Yan, D., Herrmann, S., Barnes, M. L., Barron-Gafford, G. A., Biederman, J. A., Ferrenberg, S., Fox, A. M., Hudson, A., Knowles, J. F., MacBean, N., Moore, D. J. P., Nagler, P. L., Reed, S. C., Rutherford, W. A., Scott, R. L., Wang, X., and Yang, J.: Remote sensing of dryland ecosystem structure and function: Progress, challenges, and opportunities, Remote Sens. Environ., 233, https://doi.org/10.1016/j.rse.2019.111401, 2019.

Song, C., Dannenberg, M. P., and Hwang, T.: Optical remote sensing of terrestrial ecosystem primary productivity, Prog. Phys. Geogr., 37, 834–854, https://doi.org/10.1177/0309133313507944, 2013.

Stavros, E. N., Schimel, D., Pavlick, R., Serbin, S., Swann, A., Duncanson, L., Fisher, J. B., Fassnacht, F., Ustin, S., Dubayah, R., Schweiger, A., and Wennberg, P.: ISS observations offer insights into plant function, Nat. Ecol. Evol., 1, 1–4, https://doi.org/10.1038/s41559-017-0194, 2017.

Still, C. J., Rastogi, B., Page, G. F. M., Griffith, D. M., Sibley, A., Schulze, M., Hawkins, L., Pau, S., Detto, M., and Helliker, B. R.: Imaging canopy temperature: shedding (thermal) light on ecosystem processes, New Phytol., 230, 1746–1753, https://doi.org/10.1111/nph.17321, 2021.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Peñuelas, J., and Seneviratne, S. I.: Quantifying soil moisture impacts on light use efficiency across biomes, New Phytol., 218, 1430–1449, https://doi.org/10.1111/nph.15123, 2018.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., and Peñuelas, J.: Drought impacts on terrestrial primary production underestimated by satellite monitoring, Nat. Geosci., 12, 264–270, https://doi.org/10.1038/s41561-019-0318-6, 2019.

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., and Magney, T.: Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP, Remote Sens. Environ., 209, 808–823, https://doi.org/10.1016/j.rse.2018.02.016, 2018.

Tucker, C. J.: Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ., 8, 127–150, 1979.

Turner, D. P., Gower, S. T., Cohen, W. B., Gregory, M., and Maiersperger, T. K.: Effects of spatial variability in light use efficiency on satellite-based NPP monitoring, Remote Sens. Environ., 80, 397–405, https://doi.org/10.1016/S0034-4257(01)00319-4, 2002.

Turner, D. P., Ritts, W. D., Cohen, W. B., Maeirsperger, T. K., Gower, S. T., Kirschbaum, A. A., Running, S. W., Zhao, M., Wofsy, S. C., Dunn, A. L., Law, B. E., Campbell, J. L., Oechel, W. C., Kwon, H. J., Meyers, T. P., Small, E. E., Kurc, S. A., and Gamon, J. A.: Site-level evaluation of satellite-based global terrestrial gross primary production and net primary production monitoring, Glob. Chang. Biol., 11, 666–684, https://doi.org/10.1111/j.1365-2486.2005.00936.x, 2005.

Turner, D. P., Ritts, W. D., Zhao, M., Kurc, S. A., Dunn, A. L., Wofsy, S. C., Small, E. E., and Running, S. W.: Assessing interannual variation in MODIS-based estimates of gross primary production, IEEE Trans. Geosci. Remote Sens., 44, 1899–1907, 2006a.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R., and Ahl, D. E.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ., 102, 282–292, https://doi.org/10.1016/j.rse.2006.02.017, 2006b.

Viovy, N., Arino, O., and Belward, A. S.: The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series, Int. J. Remote Sens., 13, 1585–1590, 1992.

Wan, Z.: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, Remote Sens. Environ., 140, 36–45, https://doi.org/10.1016/j.rse.2013.08.027, 2014.

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, IEEE Trans. Geosci. Remote Sens., 34, 892–905, https://doi.org/10.1109/36.508406, 1996.

Wang, X., Biederman, J. A., Knowles, J. F., Scott, R. L., Turner, A. J., Dannenberg, M. P., Köhler, P., Frankenberg, C., Litvak, M. E., Flerchinger, G. N., Law, B. E., Kwon, H., Reed, S. C., Parton, W. J., Barron-Gafford, G. A., and Smith, W. K.: Satellite solar-induced chlorophyll fluorescence and near-infrared reflectance capture complementary aspects of dryland vegetation productivity dynamics, Remote Sens. Environ., 270, 112858, https://doi.org/10.1016/j.rse.2021.112858, 2022.

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

Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., Menzer, O., and Reichstein, M.: Basic and extensible post-processing of eddy covariance flux data with REddyProc, Biogeosciences, 15, 5015–5030, https://doi.org/10.5194/bg-15-5015-2018, 2018.

Wutzler, T., Reichstein, M., Moffat, A. M., Menzer, O., Migliavacca, M., Sickel, K., and Šigut, L.: REddyProc: Post processing of (half-)hourly eddy-covariance measurements, https://cran.r-project.org/web/packages/REddyProc, 2020.

Xiao, X., Zhang, Q., Braswell, B., Urbanski, S., Boles, S., Wofsky, S., Moore, III, B., and Ojima, D.: Modeling gross primary

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

production of temperate deciduous broadleaf forest using satellite images and climate data, Remote Sens. Environ., 91, 256–270, https://doi.org/10.1016/j.rse.2004.03.010, 2004.

Yan, D., Scott, R. L., Moore, D. J. P., Biederman, J. A., and Smith, W. K.: Understanding the relationship between vegetation greenness and productivity across dryland ecosystems through the integration of PhenoCam, satellite, and eddy covariance data, Remote Sens. Environ., 223, 50–62, https://doi.org/10.1016/j.rse.2018.12.029, 2019.

Zhang, Y., Song, C., Sun, G., Band, L. E., McNulty, S., Noormets, A., Zhang, Q., and Zhang, Z.: Development of a coupled carbon and water model for estimating global gross primary productivity and evapotranspiration based on eddy flux and remote sensing data, Agric. For. Meteorol., 223, 116–131, https://doi.org/10.1016/j.agrformet.2016.04.003, 2016.

Zhang, Y., Joiner, J., Alemohammad, S. H., Zhou, S., and Gentine, P.: A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks, Biogeosciences, 15, 5779–5800, https://doi.org/10.5194/bg-15-5779-2018, 2018.

Zhang, Y., Gentine, P., Luo, X., Lian, X., Liu, Y., Zhou, S., Michalak, A. M., Sun, W., Fisher, J. B., Piao, S., and Keenan, T. F.: Increasing sensitivity of dryland vegetation greenness to precipitation due to rising atmospheric CO2, Nat. Commun., 13, 4875, https://doi.org/10.1038/s41467-022-32631-3, 2022.

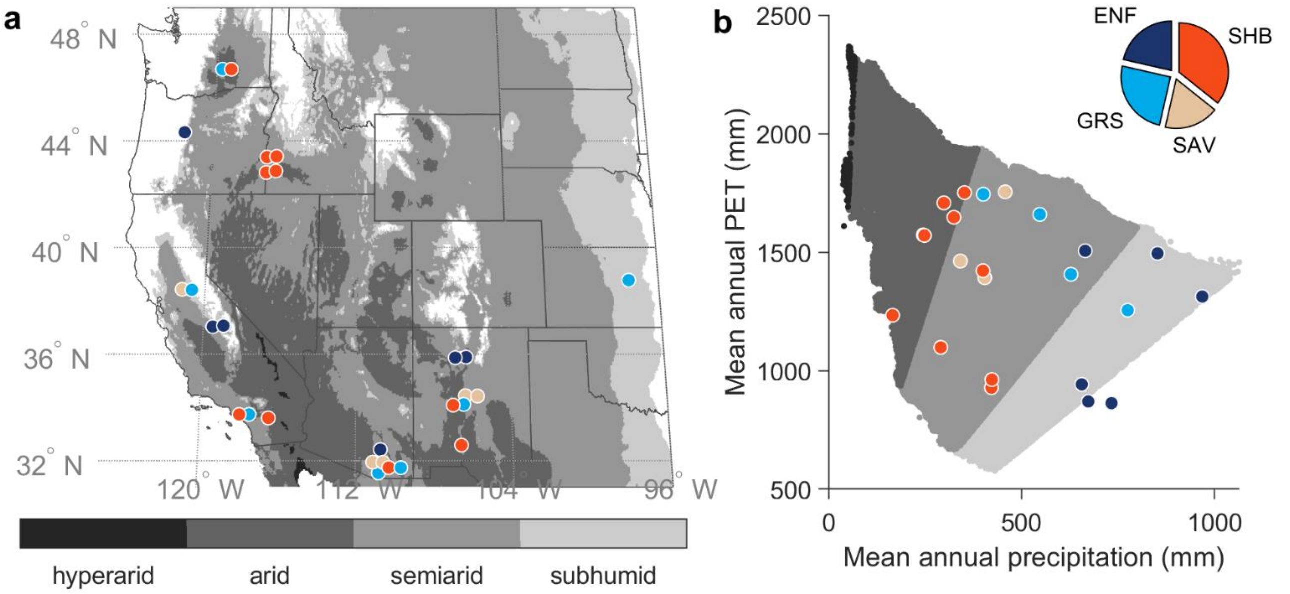
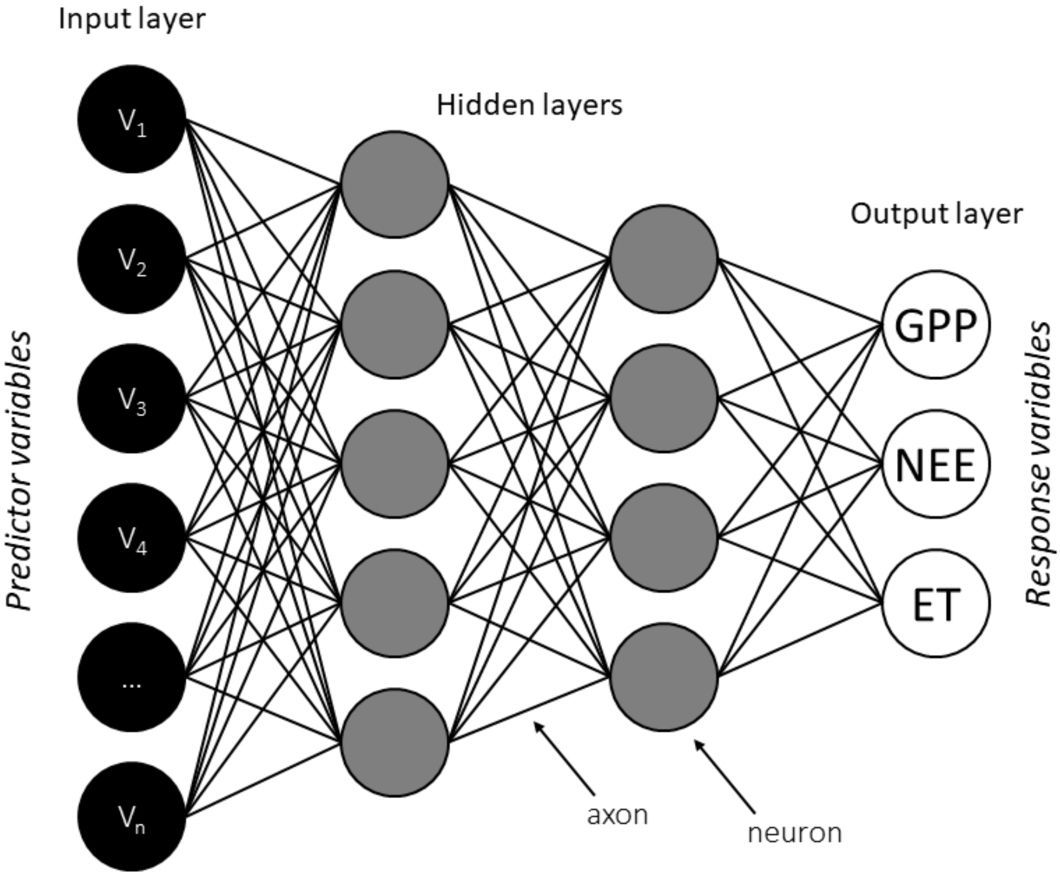


Figure 1: (a) Map of the geographic domain, aridity, and eddy covariance sites. Adjacent sites were slightly offset to improve visibility. (b) Climatic locations of the eddy covariance sites by mean annual precipitation (P) and potential evapotranspiration (PET). Gray-scale background shows the P and PET distribution of all dryland grid cells shown in (a), while the colored points show the distribution of the AmeriFlux sites in aridity space. (Note that some sites are not visible in (b) because their climates are essentially identical to adjacent sites and their points therefore perfectly overlap.) Aridity classes were defined based on the ratio of mean annual precipitation to potential evapotranspiration (over 1981-2010) from the ~4 km TerraClimate dataset (Abatzoglou et al., 2018): subhumid ($0.5 \leq P/PET \leq 0.75$), semiarid ($0.2 \leq P/PET < 0.5$), arid ($0.03 \leq P/PET < 0.2$), and hyperarid ($P/PET < 0.03$). Sites are color coded by vegetation type (shown in the pie chart in (b)): evergreen needleleaf forest (ENF), grassland (GRS), shrubland (SHB), and savanna (SAV). Site details can be found in Table S1.



775 **Figure 2: Artificial neural network (ANN) architecture conceptual diagram (Olden et al., 2008). Each "neuron" in a given layer represents a weighted combination of the neurons in the previous layer.**

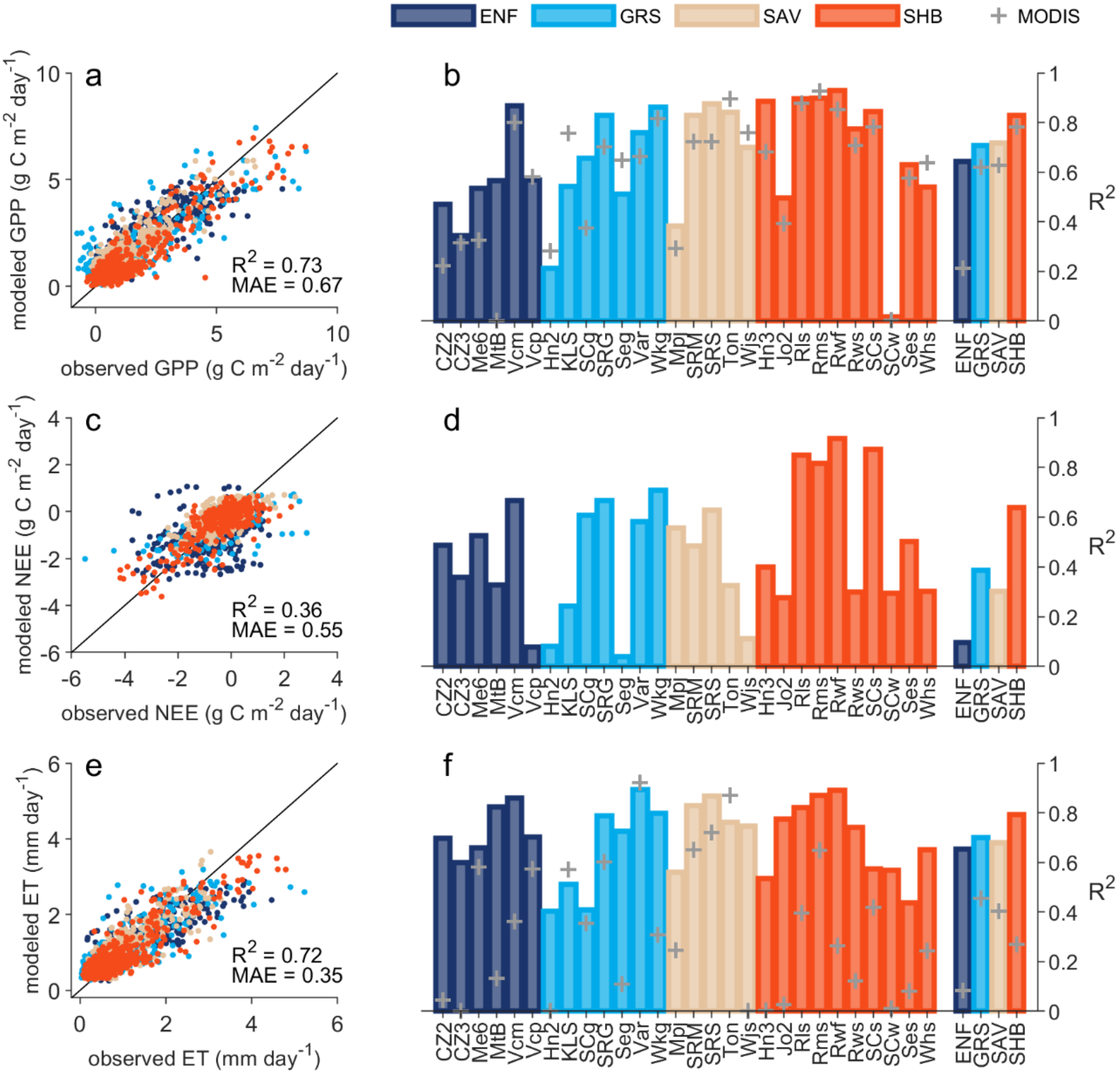


Figure 3: Overall performance of DrylANNd GPP (a-b), NEE (c-d), and ET (e-f) estimates at monthly time scale. Scatterplots show comparison of monthly tower observations to cross-validated model estimates aggregated across all 28 eddy covariance sites (Table S1). Bar plots show the R^2 between monthly tower observations and cross-validated model estimates for each site individually and aggregated by vegetation type (Fig. 1). Gray "+" signs (b & f) show the MODIS product skill for each site and vegetation type.

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

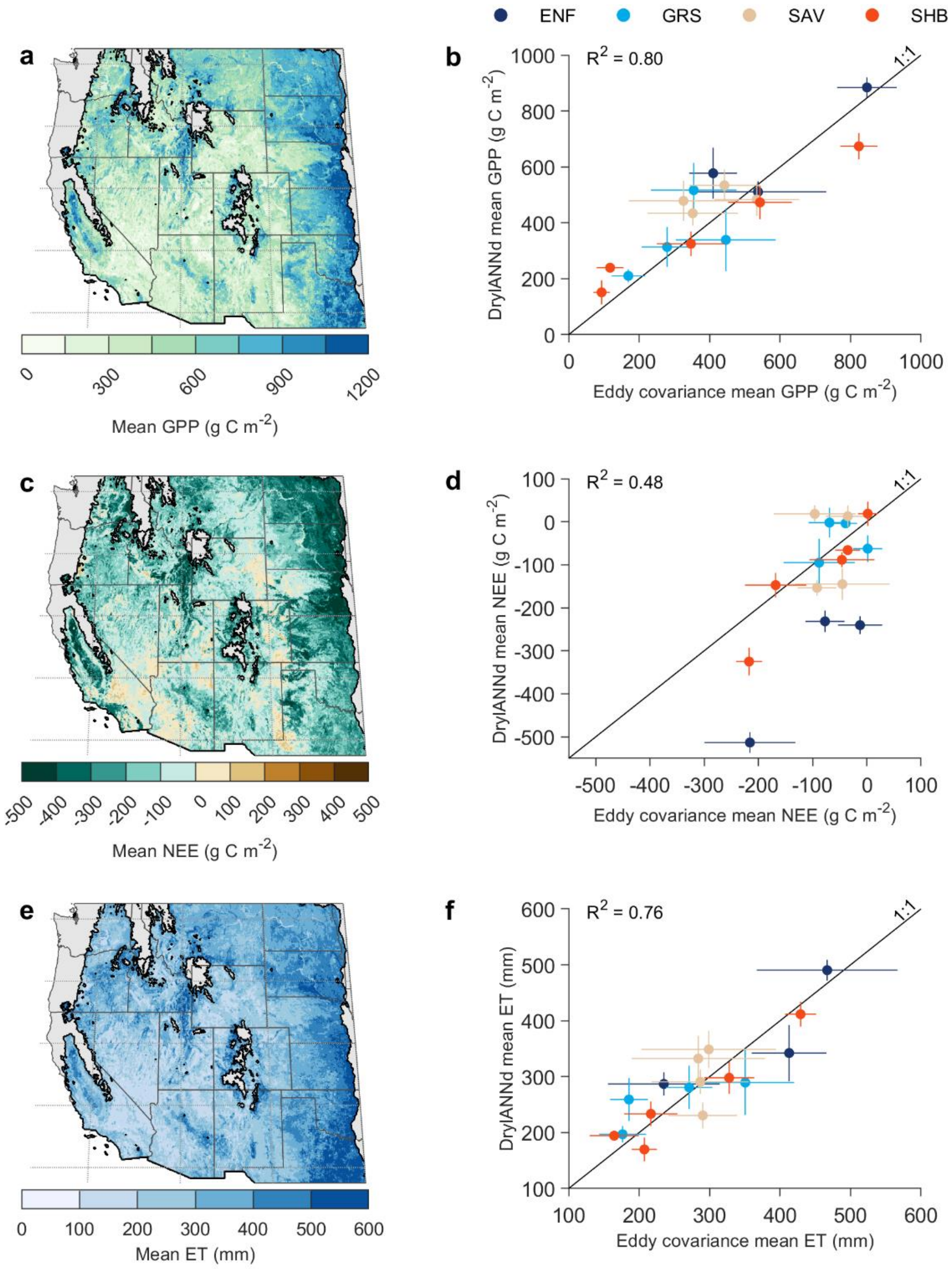


Figure 4: Spatial patterns and spatial validation of mean warm-season (April-October) GPP (a-b), NEE (c-d), and ET (e-f) predicted by DrylANNd during 2015-2020. Maps show the regional DrylANNd predictions, with the black outline indicating the region classified as dryland (P/PET ≤ 0.75). Scatterplots show mean (±1 standard deviation) warm-season fluxes predicted by DrylANNd and measured by eddy covariance at the 16 AmeriFlux sites that had complete records from 2015-2020, with points colored by vegetation type. For purposes of calculating annual means and standard errors, missing monthly data were filled with a spline interpolation.

https://doi.org/10.5194/bg-2022-186
Preprint. Discussion started: 30 September 2022

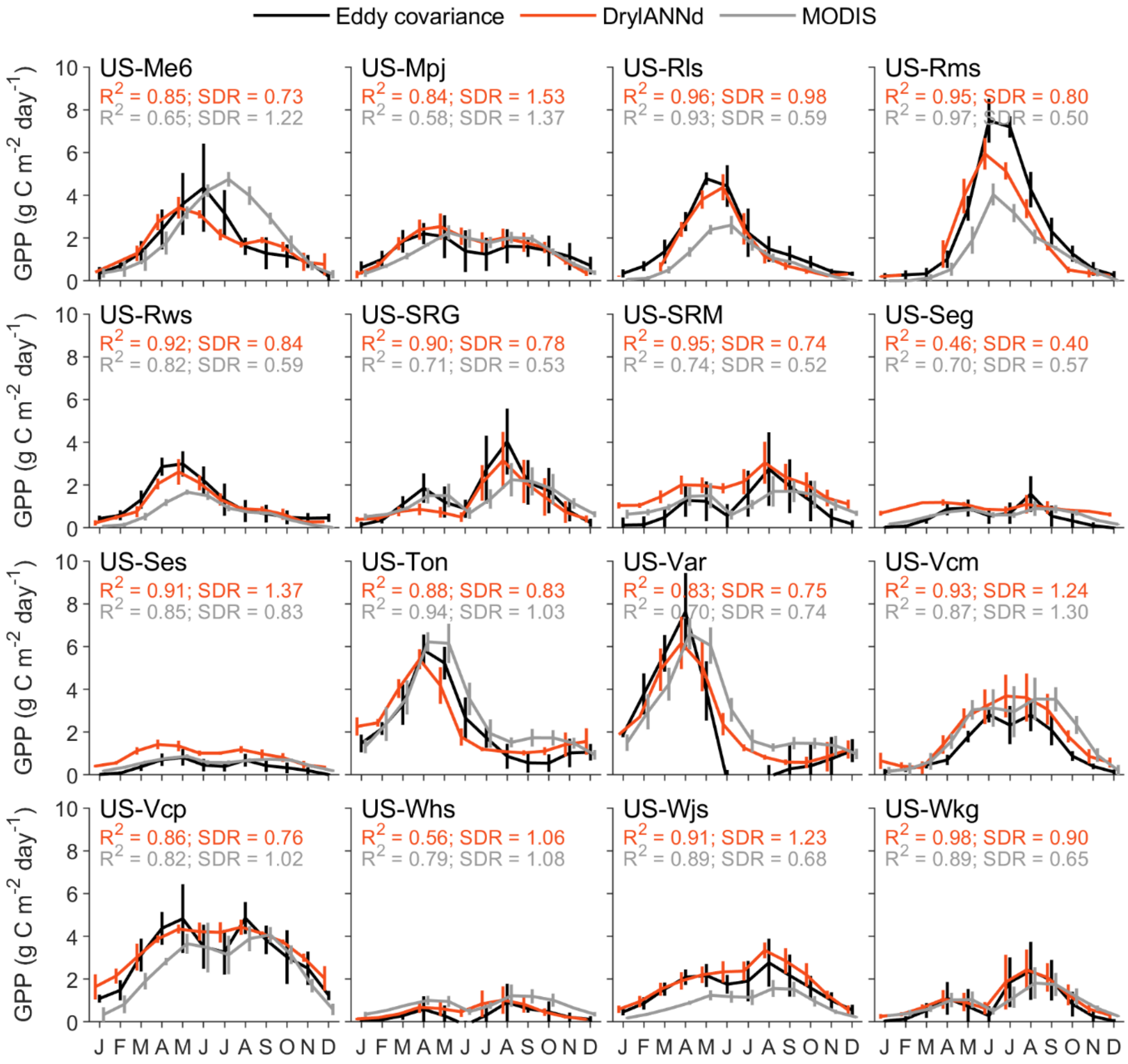


Figure 5: Mean (±1 standard deviation) seasonality of eddy covariance GPP (black), DrylANNd GPP (red), and MODIS GPP (gray) at the sixteen sites with complete flux records. Monthly means and standard deviations for all three datasets were estimated using all available observations during the period 2015-2020. Standard deviation ratio (SDR)<1 indicates that the model underestimated the magnitude of seasonal variability and SDR>1 indicates that the model overestimated the magnitude of seasonal variability.

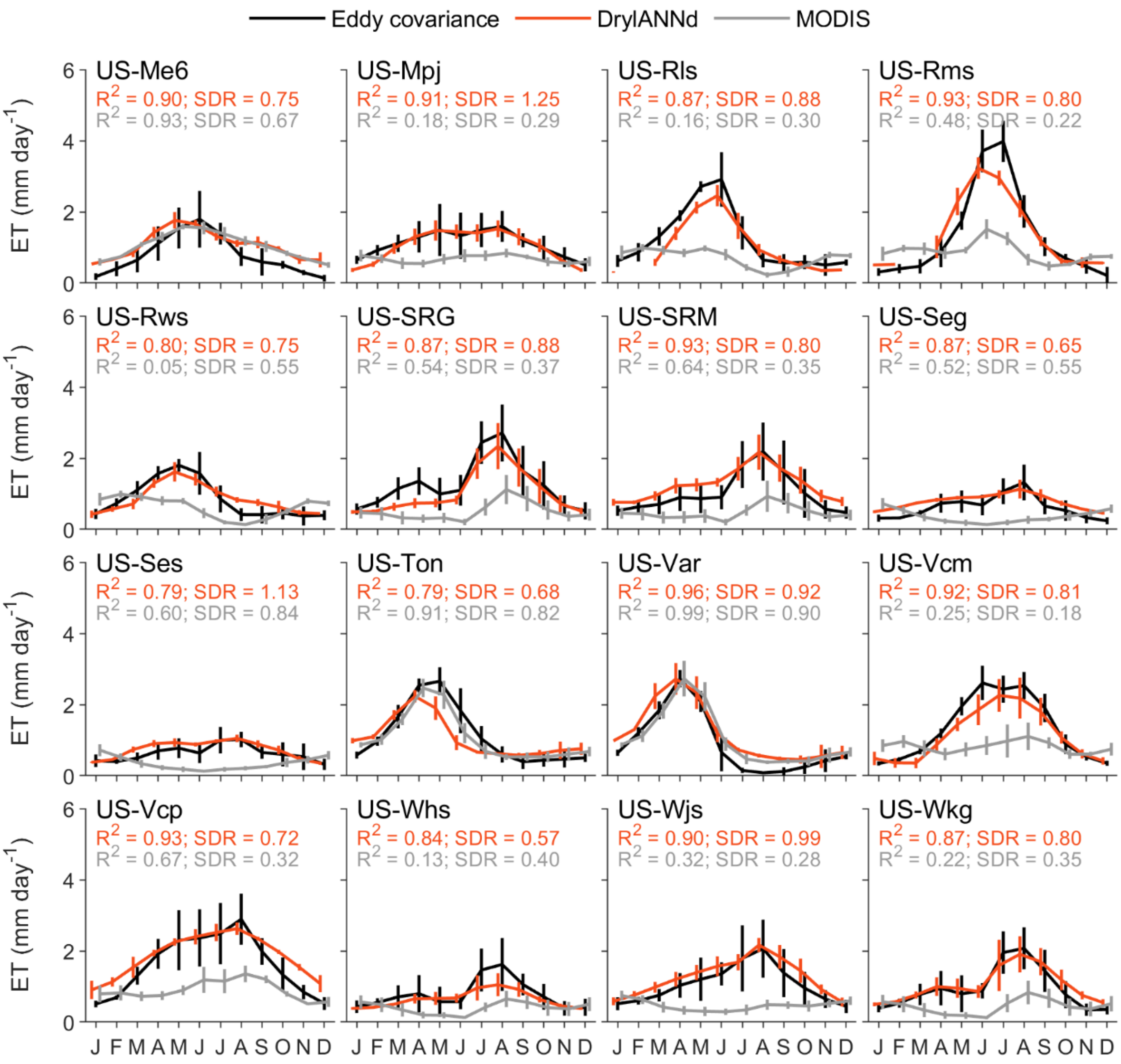
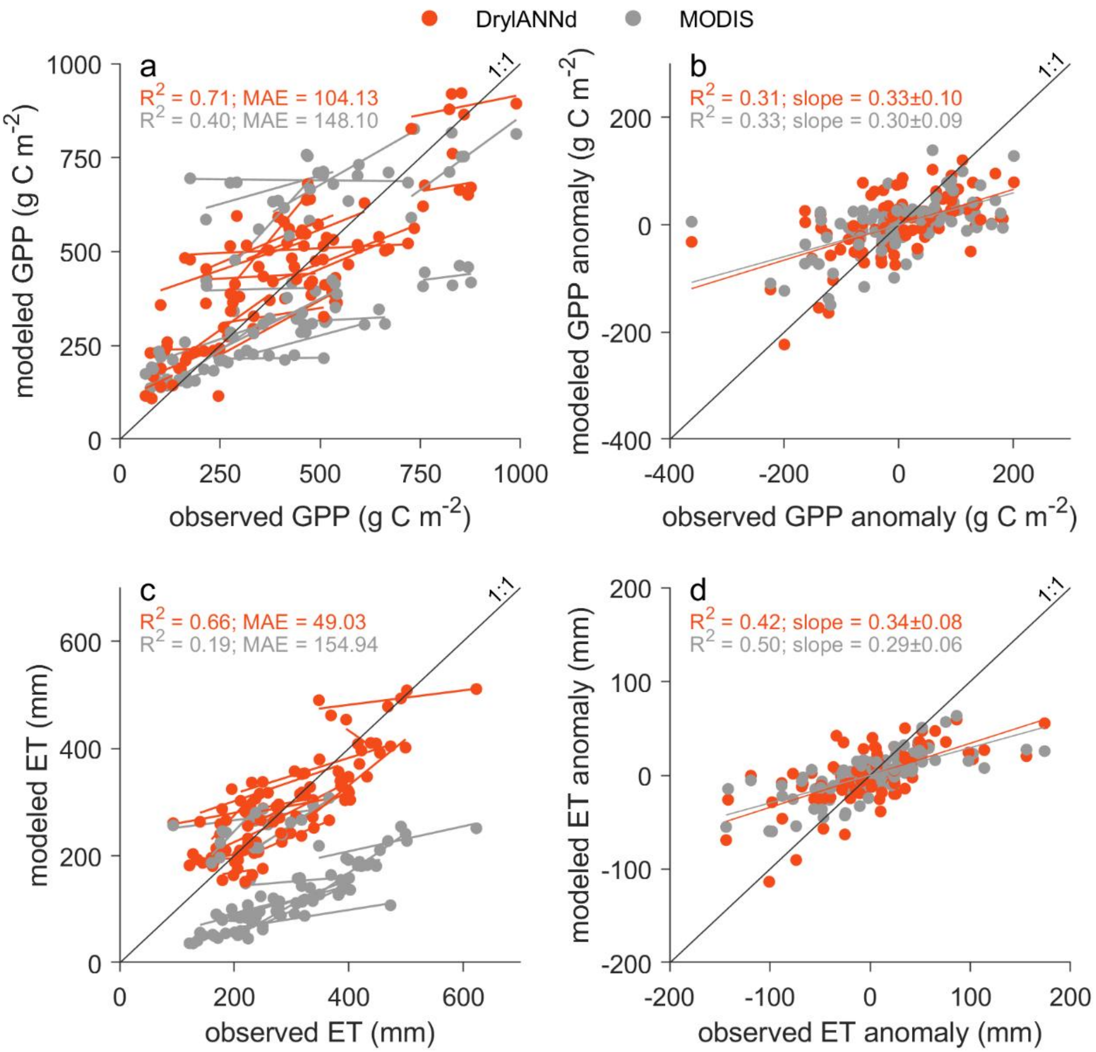


Figure 6: Mean (±1 standard deviation) seasonality of eddy covariance ET (black), DrylANNd ET (red), and MODIS ET (gray) at 800 the sixteen sites with complete flux records. Monthly means and standard deviations for all three datasets were estimated using all available observations during the period 2015-2020. Standard deviation ratio (SDR)<1 indicates that the model underestimated the magnitude of seasonal variability and SDR>1 indicates that the model overestimated the magnitude of seasonal variability.



805 **Figure 7: Annual and interannual variability of warm season (April-October) (a-b) GPP and (c-d) ET. Lines in (a) and (c) show the linear relationship between estimated and observed GPP and ET for each site, while lines in (b) and (d) show the linear relationship between estimate and observed GPP and ET anomalies across all sites. For purposes of calculating annual values and anomalies, missing monthly data were filled with a spline interpolation.**

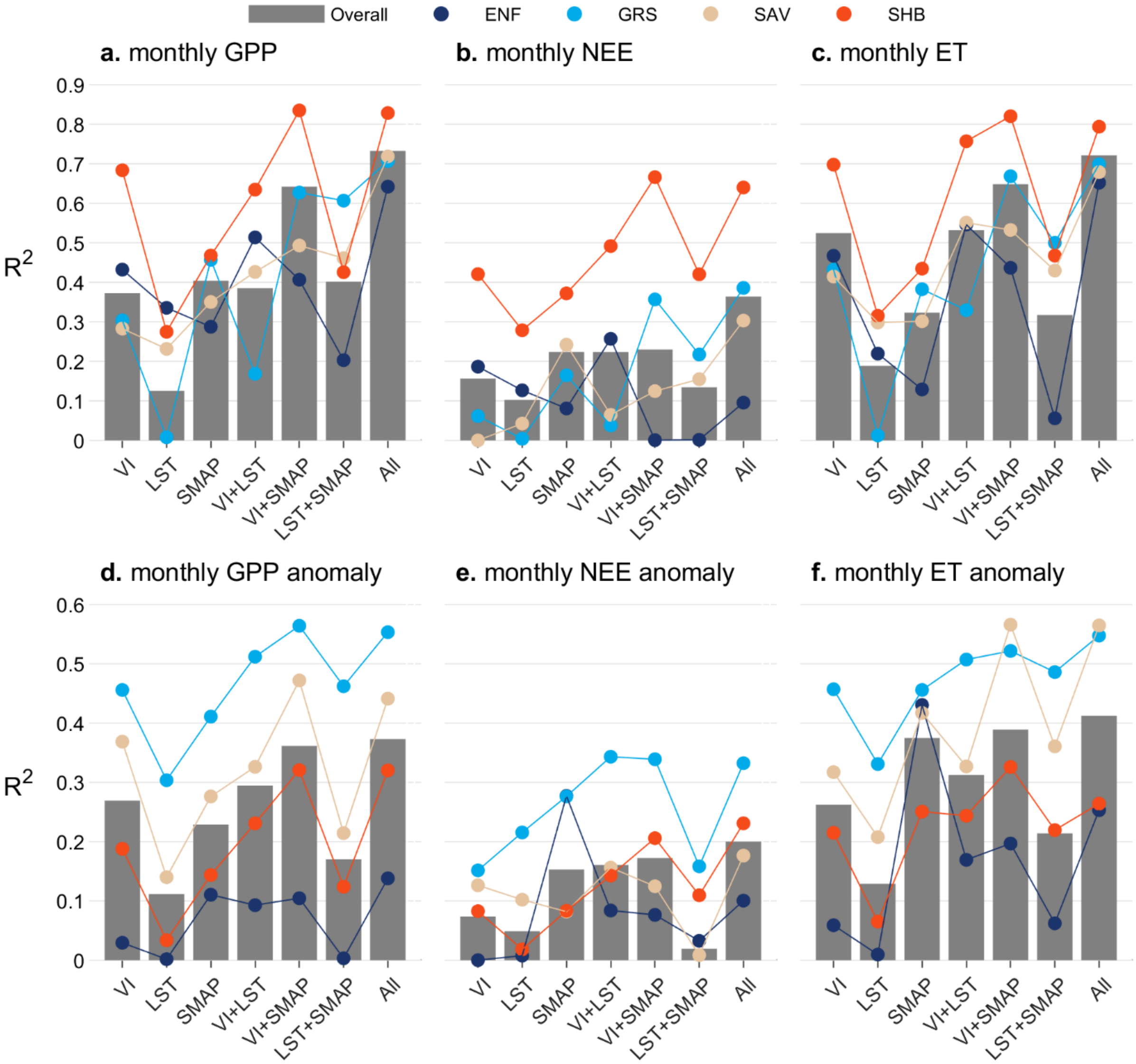
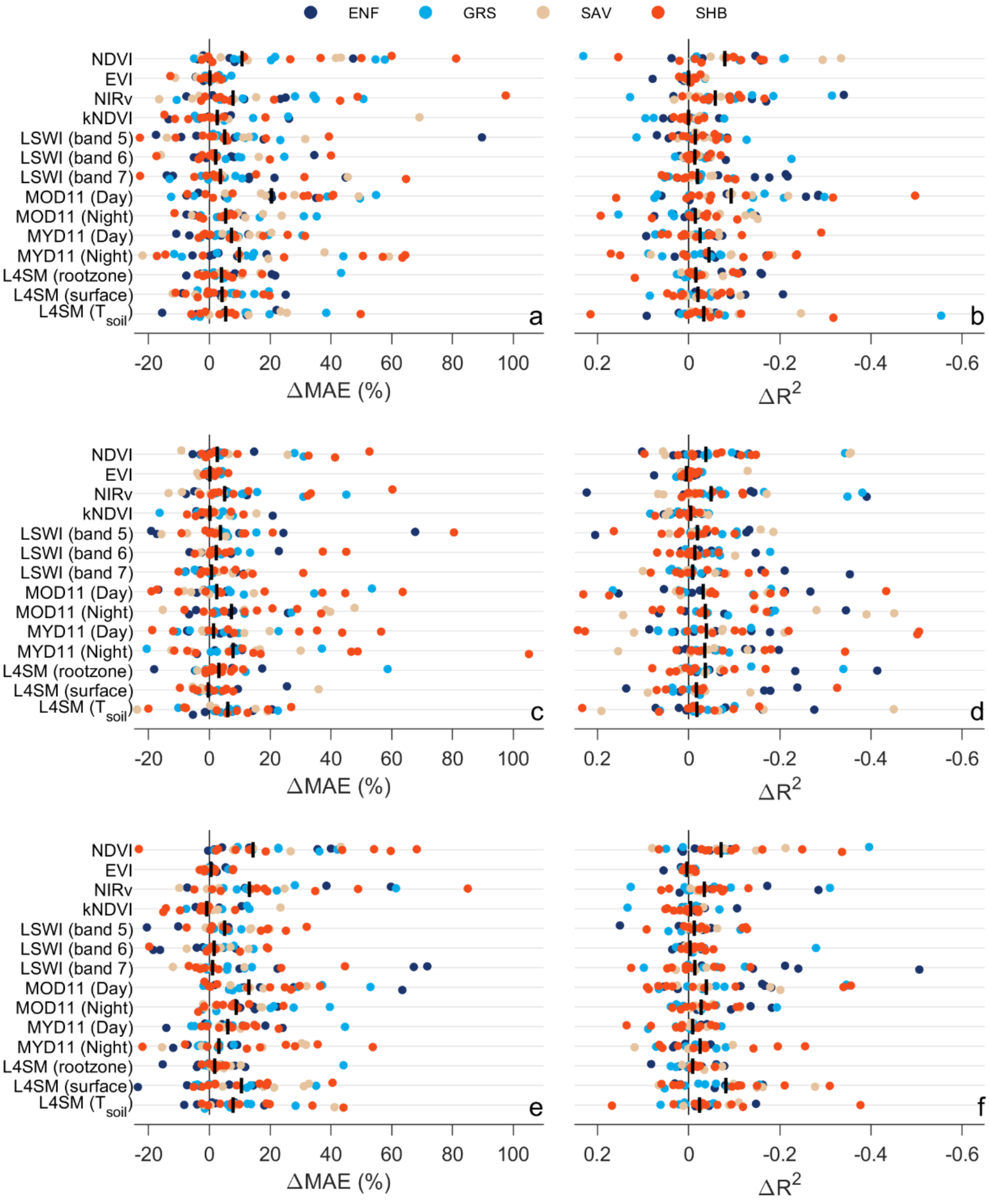


Figure 8: Coefficient of determination (R^2) for models based on different combinations of predictor variables: optical VIs, MODIS LST, and SMAP soil moisture/temperature. Gray bars show overall model R^2 (across all sites) while colored dots show model R^2 across all sites of a given vegetation type. (a-c) DrylANNd model performance for predicting observed monthly fluxes across all 28 eddy covariance sites. (d-f) DrylANNd model performance for predicting monthly flux anomalies relative to monthly site-means across the 16 eddy covariance sites that cover the full study period (2015-2020). For purposes of calculating monthly site-means, missing monthly data were filled with a spline interpolation after which mean fluxes were calculated for each month during the 2015-2020 study period.



820 **Figure 9: Site-level (filled circles, colored by vegetation type) and median (black vertical line) change in MAE (ΔMAE) and R^2 (ΔR^2) for DrylANNd model predictions of (a-b) GPP, (c-d) NEE, and (e-f) ET when the information content of each variable is destroyed via random permutation. For each site, 100 random permutations of each variable were performed (each with the same mean and variance as the original variable), model predictions were generated with a randomly-selected ANN (from which that site was excluded from model calibration) from the ensemble for each noise-only permutation, and predictions with the noise-only simulation**
825 **was compared to those with the original variable.**

Table 1. Remotely sensed predictor variables used in the DrylANNd carbon and water flux model.

Product (site / region)	Variable	Units	Resolution (site/region)
MODIS NBAR (MCD43A4 / MCD43C4)	NDVI	-	500 m / 0.05°
	EVI	-	500 m / 0.05°
	NIRv	-	500 m / 0.05°
	kNDVI	-	500 m / 0.05°
	LSWI (1240 nm)	-	500 m / 0.05°
	LSWI (1640 nm)	-	500 m / 0.05°
	LSWI (2130 nm)	-	500 m / 0.05°
MODIS Terra LST (MOD11A1 / MOD11C1)	Day LST (~10:30 a.m.)	K	1 km / 0.05°
	Night LST (~10:30 p.m.)	K	1 km / 0.05°
MODIS Aqua LST (MYD11A1 / MYD11C1)	Day LST (~2:30 p.m.)	K	1 km / 0.05°
	Night LST (~2:30 a.m.)	K	1 km / 0.05°
SMAP Level 4 Soil Moisture (SMAP L4SM)	Rootzone (0-100 cm) soil moisture	%	9 km[a]
	Surface (0-5 cm) soil moisture	%	9 km[a]
	Soil temperature	K	9 km[a]
Rangeland Analysis V3 Fractional Cover	Annual grasses & forbs	%	30 m[b]
	Perennial grasses & forbs	%	30 m[b]
	Tree	%	30 m[b]
	Shrub	%	30 m[b]
	Bare ground	%	30 m[b]
	Litter	%	30 m[b]

[a] Resampled to 0.05° using nearest neighbor for regional scale.

[b] Averaged within 500 m buffer for site-level calibration and within 0.05° grid for regional scaling.