CO₂ exchange and evapotranspiration across dryland ecosystems of southwestern North America


¹Southwest Watershed Research Center, Agricultural Research Service, Tucson, AZ, USA
²Earth Research Institute, University of California Santa Barbara, Santa Barbara, CA, USA
³Department of Biology, University of Utah, Salt Lake City, UT, USA
⁴School of Forestry, Merriam-Powell Center for Environmental Research, Northern Arizona University, Flagstaff, AZ, USA
⁵Departamento de Ciencias del Agua y Medio Ambiente, Instituto Tecnológico de Sonora, Ciudad Obregón, Sonora, Mexico
⁶Atmospheric Turbulence and Diffusion Division, Air Resources Laboratory, National Oceanographic and Atmospheric Administration, Oak Ridge, TN, USA
⁷Department of Biology, University of New Mexico, Albuquerque, NM, USA
⁸Global Change Research Group, Department of Biology, San Diego State University, San Diego, CA, USA
⁹Department of Geography, College of Life and Environmental Sciences, Exeter, UK
¹⁰School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA
¹¹Departamento de Agricultura y Ganaderia, Universidad de Sonora, Hermosillo, Sonora, Mexico
¹²Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA
¹³Departamento de Fisica, Universidad de Sonora, Hermosillo, Sonora, Mexico
¹⁴Department of Earth System Science, University of California Irvine, Irvine, CA, USA

Correspondence
Joel Biederman, USDA-ARS, Tucson, AZ, USA.
Email: joel.biederman@ars.usda.gov

Funding information
U.S. Department of Energy’s Office of Science; National Science Foundation, Grant/Award Number: EAR-125501; NSF SGER; SDSU; SDSU Field Stations; NSF International Program; CONACYT; CIBNOR

Abstract
Global-scale studies suggest that dryland ecosystems dominate an increasing trend in the magnitude and interannual variability of the land CO₂ sink. However, such analyses are poorly constrained by measured CO₂ exchange in drylands. Here we address this observation gap with eddy covariance data from 25 sites in the water-limited Southwest region of North America with observed ranges in annual precipitation of 100–1000 mm, annual temperatures of 2–25°C, and records of 3–10 years (150 site-years in total). Annual fluxes were integrated using site-specific ecohydrologic years to group precipitation with resulting ecosystem exchanges. We found a wide range of carbon sink/source function, with mean annual net ecosystem production (NEP) varying from -350 to +330 gCm⁻² across sites with diverse vegetation types, contrasting with the more constant sink typically measured in mesic ecosystems. In this region, only forest-dominated sites were consistent carbon sinks. Interannual variability of NEP, gross ecosystem production (GEP), and ecosystem respiration (Reco) was larger.
than for mesic regions, and half the sites switched between functioning as C sinks/C sources in wet/dry years. The sites demonstrated coherent responses of GEP and NEP to anomalies in annual evapotranspiration (ET), used here as a proxy for annually available water after hydrologic losses. Notably, GEP and $R_{eco}$ were negatively related to temperature, both interannually within site and spatially across sites, in contrast to positive temperature effects commonly reported for mesic ecosystems. Models based on MODIS satellite observations matched the cross-site spatial pattern in mean annual GEP but consistently underestimated mean annual ET by ~50%. Importantly, the MODIS-based models captured only 20–30% of interannual variation magnitude. These results suggest the contribution of this dryland region to variability of regional to global CO$_2$ exchange may be up to 3–5 times larger than current estimates.

**KEYWORDS**
climate, moderate resolution imaging spectroradiometer, net ecosystem exchange, photosynthesis, remote sensing, respiration, semiarid, water

1 | INTRODUCTION

Dryland ecosystems (arid and semiarid) occupy approximately 40% of the terrestrial surface (Reynolds et al., 2007), and global-scale studies suggest they strongly influence the carbon cycle due to the inherent variability of water, their main limiting resource (Ahlström et al., 2015; Jung et al., 2011; Middleton & Thomas, 1992; Poulter et al., 2014). Unfortunately, the availability of continuous, long-term measurements of water and CO$_2$ exchange has lagged in drylands as compared to wetter regions (mesic and humid). Therefore, understanding of dryland ecosystem-atmosphere exchanges relies heavily upon remote sensing models (RSM), coarse-scale atmospheric inversion models, empirical regression models, or terrestrial biosphere models, all of which are poorly constrained by dryland ecosystem data (Ahlström et al., 2015; Jung et al., 2011; Ma, Huete, Moran, Ponce-Campos, & Eamus, 2015; Poulter et al., 2014; Verma et al., 2014; Xiao et al., 2014). While much has been learned from field measurements such as plot-scale aboveground net primary production (ANPP) (Huxman et al., 2004; Lauenroth & Sala, 1992; Noy-Meir, 1973; Ponce-Campos et al., 2013), the maturation of multi-annual, multisite eddy covariance datasets presents an opportunity to advance understanding of (i) diverse dryland ecosystems for which ANPP data lack; (ii) tower footprint-scale net and gross CO$_2$ fluxes from entire ecosystems (i.e., above- and below-ground); and (iii) bidirectional, coupled, land-atmosphere exchanges of CO$_2$ and water, the critical limiting resource in drylands.

Until recently, efforts to synthesize direct observations of CO$_2$ and water exchange across multiple eddy covariance sites were weighted toward productive, mesic ecosystems, especially forests (Baldocchi, 2008; Chen et al., 2015; Keenan et al., 2012; Law et al., 2002; Luyssaert et al., 2007; Valentini et al., 2000; Williams et al., 2012). These studies suggest that annual NEP is usually positive (sink of ~100–300 gCm$^{-2}$) and that annual GEP is positively related to both water availability and temperature (Yi et al., 2010), while annual ET tends to closely track seasonal changes in available energy (Williams et al., 2012). Meanwhile, annual $R_{eco}$ appears to be positively related to mean annual temperature in some cases (Reichstein et al., 2007; Yu et al., 2013) or unrelated in others (Janssens et al., 2001). In mesic to humid systems, where water is not limiting, annual ET tends to be relatively dependent on available energy (Williams et al., 2012). Drought tends to reduce GEP more strongly than $R_{eco}$, resulting in reduced, but usually positive, net CO$_2$ uptake (Baldocchi, 2008; Reichstein et al., 2007; Schwalm et al., 2010). Flux data syntheses weighted to mesic sites suggest relative interannual variability of GEP and $R_{eco}$ ranges from ~5 to 20% (Keenan et al., 2012; Yuan et al., 2009), whereas less is known about interannual variability of ET, especially for drylands (Villarreal et al., 2016). While model studies suggest larger interannual variability in dryland fluxes (Ahlström et al., 2015; Jung et al., 2011), measured interannual variability remains largely unknown (Tramontana et al., 2016).

While a growing number of studies present dryland fluxes, none has included sufficient sites and years of data to characterize the regional magnitude and interannual variability of CO$_2$ exchange and ET (Anderson-Teixeira, Delong, Fox, Brese, & Litvak, 2011; Baldocchi, 2014; Dore et al., 2010; Luo et al., 2007; Ma, Baldocchi, Xu, & Hehn, 2007; Pereira et al., 2007; Scott, Biederman, Hamerlynck, & Barron-Gafford, 2015; Thomas et al., 2009). These studies suggest unique dryland characteristics requiring more comprehensive evaluation including: (i) a wider range of site mean NEP than mesic biomes, including persistent carbon sources and sinks (Baldocchi, 2008; Biederman et al., 2016; Ma, Baldocchi, Wolf, & Verfaillie, 2016; Yu et al., 2013); (ii) large interannual variability of CO$_2$ exchange and ET, including sites that 'pivot' between carbon sink/source function at site-specific 'pivot points' in water availability, demonstrating the strong linkage between CO$_2$ and water in drylands (Kurc & Small, 2007; Ma et al., 2016; Pereira et al., 2007; Scott et al., 2015); and (iii) unique seasonal dynamics of temperature, precipitation,
phenology, which can sometimes drive two distinct seasons (winter/ spring, summer) of elevated CO₂ exchange and ET, depending on the relative timing of water and radiation inputs (Anderson-Teixeira et al., 2011; Dore et al., 2010; Scott, Serrano-Ortiz, Domingo, Hamerlynck, & Kowalski, 2012). These unique traits call for caution in extrapolating ecological understanding from measurements in more well-studied mesic regions and highlight the need for directly comparable measurements in drylands.

In the absence of widespread ground observations, RSM are commonly used to estimate dryland CO₂ exchange and ET (Ma et al., 2015; Ponce-Campos et al., 2013; Poulter et al., 2014). Satellite-derived greenness (EVI, NDVI) may be empirically related to point measurements of CO₂ exchange and ET (Goulden et al., 2012; Ponce-Campos et al., 2013; Sims et al., 2006) or used with climate data to derive process-based RSM for GEP and ET (Mu, Zhao, & Running, 2011; Sims et al., 2006; Smith et al., 2016; Turner et al., 2006). Evaluation of RSM across flux tower networks suggests that RSM capture spatial patterns of GEP but contain relatively little information about GEP interannual variability (Jin & Goulden, 2014; Verma et al., 2014). RSM are expected to find particular challenges in drylands due mainly to (i) satellite-derived greenness indices that do not capture the physiological properties of the vegetation that control photosynthesis, and (ii) assumptions of static, biome-level parameters (e.g., light-use efficiency), which vary seasonally with root zone soil moisture and other factors (Krofcheck et al., 2015; Kurc & Small, 2007; Liu, Rambal, & Mouillot, 2015; Verma et al., 2014). Prior data-RSM comparisons have included few dryland sites (Mu et al., 2011; Sims et al., 2006; Turner et al., 2006), and there is a need to understand how well commonly used RSM capture the magnitude and interannual variability of measured CO₂ exchange and ET.

Here we address the gap in dryland flux observations and evaluate whether common assumptions from mesic regions are valid in a water-limited region. We present a synthesis of site-level climate and ecosystem-atmosphere exchange of water and CO₂ at monthly and annual time scales for 25 sites with a combined 150 site-years of measurements. The questions addressed here are: (i) Do the spatially diverse seasonal patterns of precipitation, temperature, and phenology combine to produce identifiable patterns of CO₂ exchange and ET across the Southwest region of North America? We expect more diverse seasonality than in mesic regions including winter-dominated, summer-dominated, and bimodal ecosystems. (ii) What are the magnitudes of spatial and interannual variability CO₂ exchange and ET? We expect variability to be larger in water-limited ecosystems than in mesic ecosystems limited by energy or nutrients (Ma et al., 2015; Noy-Meir, 1973; Yuan et al., 2009). (iii) How well do MODIS-based models capture spatial and interannual variability of measured CO₂ exchange and ET? Detailed diagnosis of RSM errors is beyond the scope of this paper, as our purpose here is to provide a first model-data comparison focused on the magnitude and interannual variability of CO₂ exchange and ET fluxes in this water-limited region.

2 MATERIALS AND METHODS

2.1 Study region: southwestern North America

The study region (hereafter referred to as ‘the Southwest’) comprises sites in the US states of Arizona, New Mexico, southern portions of Utah and California, and the Mexican states of Sonora and Baja California Sur (Figure 1a). The Southwest is characterized by water limitation at the annual scale (potential ET >> precipitation). The Southwest has large spatial gradients in mean annual precipitation (MAP 40–1200 mm, Figure 1) and temperature (MAT –4 to +26°C, Fig. S1) due to interactions among topography, latitude, wind patterns, and distance from oceans. The 25 study sites represent the majority of the Southwest in this two-dimensional climate space, especially the most frequent portion of MAT –10–20°C and MAP –200–600 mm (Figure 1b).

Seasonal weather dynamics drive distinct spatial patterns of ecosystem CO₂ exchange and ET across the Southwest. The western portion adjacent to the Pacific Ocean experiences a Mediterranean climate with up to 90% of MAP occurring during winter months (defined here as November–April, Fig. S2). In contrast, the central, eastern, and southeastern portions are heavily influenced by the North American Monsoon (Douglas, Maddox, Howard, & Reyes, 1993), which brings a majority of the annual precipitation during summer (June–October) from moisture sources in the Gulf of Mexico and California (Fig. S2). Across the Southwest, winter precipitation from Pacific frontal systems tends to be greater during negative phases of the El Niño Southern Oscillation (ENSO) and less during positive, La Niña phases (Andrade & Sellers, 1988). Baja California receives precipitation primarily from Pacific tropical storms during early fall (September–October) and is influenced by teleconnections with the Southern California current (Reimer et al., 2015).

2.2 Study sites

We used 25 eddy covariance flux sites with 3–10 years of measurements (mean of 6 years, total n = 150 years) representing the climate and ecosystems of the Southwest (Figure 1, Table 1, Table S1, Figs S1 and S2). Observations were made between 1999 and 2014, with greatest data density during 2005–2014 (Table 1). The major regional IGBP vegetation classes represented, identified by site teams, include grassland, open and closed shrubland, savanna and woody savanna, mixed forest, and evergreen needleleaf forest. Study sites were initially classified into based on based on Köppen climate (http://koeppen-geiger.vu-wien.ac.at/present.htm). However, patterns of distinct climate and flux dynamics (Figure 2) warranted some separation/combination of Köppen classes, resulting in seven subregional groups (Table 1, Figure 1). For site details, see references in Table 1 and www.fluxdata.org.

2.3 Flux data collection and processing

We assembled 30-min measurements of precipitation, ET, NEP, and GEP and R_{eco} derived from NEP using methods similar to those
described in Biederman et al. (2016). Measurements of terrestrial-atmosphere gas exchange were made at the ecosystem level using the eddy covariance technique (Goulden, Munger, Fan, Daube, & Wofsy, 1996). Data collection and regular calibrations of eddy covariance flux measurement systems followed accepted guidelines (Lee, Massman, & Law, 2006). Periods of insufficient turbulent mixing were screened using a friction velocity filter (Reichstein et al., 2005). Gross fluxes were partitioned from the net CO2 exchange measurements based on the observed relationship between nighttime respiration and temperature, which was then used to separately derive daytime GEP and Reco.

Precipitation falling late in the calendar year is often stored as snowpack or soil moisture and should be associated with plant productivity in the following growing season (Ma et al., 2007; Pereira et al., 2007; Thomas et al., 2009). Therefore, annual flux sums and metrological driving variables were calculated using an ecohydrologic year spanning November–October. Because multiyear records at 12 of the sites commenced with the calendar year, the first two months of these records (November and December of the first hydrologic year in each record) were filled using mean monthly fluxes of remaining years (Biederman et al., 2016), representing 2–5% of data presented. This gap filling had minimal impact on annual sums because November and December are relatively low-flux months, together comprising only 5–15% of mean annual fluxes at a given site (Figure 2). No filling of missing months was applied at MX-IPA, which has warmer winters with higher fluxes.

Because US-cop is the only long-term eddy covariance dataset in the dry northern portion of the Colorado Plateau, we took additional steps to produce annual sums from the published weekly values (Bowling, Bethers-Marchetti, Lunch, Grote, & Belnap, 2010), enabling the site to be included here. Multiple linear regression was used to gap-fill missing weekly values of NEP, GEP, and ET using 10-cm soil moisture, solar radiation, and air temperature with separate regression models for each season. Reco was calculated as NEP minus GEP. Approximately 20% of weekly values at US-cop were thus modeled, with the majority of filled weeks (>70%) occurring during low-flux winter periods.

2.4 Use of ET as a proxy for ecosystem water availability

As the most widely measured hydrologic flux, precipitation (P) is the common proxy for annual ecosystem water availability (Sala,
<table>
<thead>
<tr>
<th>Subregion</th>
<th>Site ID</th>
<th>Description</th>
<th>Dominant species</th>
<th>IGBP class</th>
<th>Koppen climate</th>
<th>Elevation (m)</th>
<th>MAP (mm)</th>
<th>MAT (°C)</th>
<th>Obs. years</th>
<th>Site reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>subP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monsoon high elevation</td>
<td>US-vcm</td>
<td>Valles Caldera mixed conifer forest</td>
<td>Picea engelmanni, Picea pugens, Abies lasiocarpa var. lasiocarpa, Abies concolor</td>
<td>Evergreen needleleaf forest</td>
<td>Dfb</td>
<td>3042</td>
<td>724</td>
<td>2.9</td>
<td>2007-2012</td>
<td>Anderson-Teixeira et al. (2011)</td>
</tr>
<tr>
<td>Monsoon mid elevation</td>
<td>US-vcp</td>
<td>Valles Caldera ponderosa forest</td>
<td>Pinus ponderosa, Quercus gambeli</td>
<td>Evergreen needleleaf forest</td>
<td>Dfb</td>
<td>2501</td>
<td>547</td>
<td>5.7</td>
<td>2007-2012</td>
<td>Anderson-Teixeira et al. (2011)</td>
</tr>
<tr>
<td>Subregion</td>
<td>Site ID</td>
<td>Description</td>
<td>Dominant species</td>
<td>IGBP class</td>
<td>Koppen climate</td>
<td>Elevation (m)</td>
<td>MAP (mm)</td>
<td>MAT (°C)</td>
<td>Obs. years</td>
<td>Site reference</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>MX-ray</td>
<td>Rayon subtropical shrubland</td>
<td><em>Fouquieria macdougalii</em>, <em>Acacia cochlacanthana</em>, <em>Parkinsonia procox</em>, <em>Mimoa distachya</em>, <em>Prosopis velutina</em></td>
<td>Savannah</td>
<td>Bsh</td>
<td>634</td>
<td>475</td>
<td>21.7</td>
<td>2009-2012</td>
<td>Méndez-Barroso et al. (2014)</td>
</tr>
</tbody>
</table>

*Mean annual temperature (MAT) and mean annual precipitation (MAP) 1950–2000 from the WorldClim 1-km dataset (www.worldclim.org).*  
*Observations collected at soo after a 2003 stand-replacing wildfire are denoted ‘sob’ and analyzed as a unique site.*
Gherardi, Reichmann, Jobbágy, & Peters, 2012). However, some portion of $P$ is lost to runoff and drainage and is unavailable to the ecosystem. Therefore, we also used ET as a metric of annual water availability. At shorter time scales (diurnal to seasonal), rates of ET are related to many factors including available energy, root density, leaf area and soil water retention properties. However, water-limited ecosystems frequently experience exhaustions of soil moisture, the limiting resource, throughout the year (Thomas et al., 2009). Regardless of shorter-term temporal dynamics, annually integrated ET represents the efflux of soil moisture that has been available to drive ecosystem exchanges after some precipitation is lost to runoff and drainage ($Q$) (Biederman et al., 2016; Briggs & Shantz, 1913; Noy-Meir, 1973). The local water balance reflects this hydrologic partitioning of $P$ to hydrologic losses ($Q$), storage changes ($S$, usually negligible at the annual scale), and soil moisture recharge, the main source of ET (Equation 1).

$$P = ET + Q - S.$$ (1)

Hydrologic losses ($Q$) tend to be larger in wetter years and at wetter sites, increasing the amount by which precipitation overestimates water available to drive CO$_2$ exchange (Biederman et al., 2016; Ponce-Campos et al., 2013; Villarreal et al., 2016). While ecosystem-level ET does not provide details on the partitioning of evaporation and transpiration, removal of hydrologic losses makes ET a better metric of available water than precipitation.

### 2.5 Remote sensing model estimates of ecosystem exchange

We used the enhanced vegetation index (EVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 5. We initially used EVI to explore seasonal similarities between greenness and fluxes of CO$_2$ and water. As RSM of these fluxes rely on satellite indices closely correlated to EVI, this comparison provides insight into RSM performance across this region. This index is distributed as MOD13Q1 data product with 16-day temporal resolution. We used the single 250-m pixel containing each flux tower. Mean monthly EVI values were computed using data classified as best quality to reduced contamination associated with clouds, shadows, and snow/ice. We obtained monthly aggregates of MODIS GEP (MOD17A2) and ET (MOD16A2) for the 1-km pixel containing each flux tower.
from the Numerical Terradynamic Simulation Group (http://files.ntsg.umt.edu/data/NTSG_Products/). MODIS ET calculates daily estimates based on the Penman-Monteith equation, driven by multiple satellite-based data (FPAR, LAI, albedo) as well as daily meteorological reanalysis (Mu et al., 2011). MODIS GEP calculates daily estimates based on light-use efficiency logic, driven by multiple satellite datasets (FPAR, LAI) as well as daily meteorological reanalysis. For more details on the MODIS GEP and ET algorithms refer to Mu et al. (2011) and Zhao & Running (2010), respectively.

2.6 Calculation of annual anomalies

We assessed regional patterns in climate and flux anomalies with time series of annual z-scores (standard deviations from the mean annual value at a given site). Because the years of measurements at each site varied, we defined a common baseline around which to compute z-scores by excluding regionally wet years (2004, 2005, 2010) and dry years (2002, 2003, 2012) during which precipitation was more than one standard deviation from the mean (1999–2014) according to the NOAA climate data for the US Southwest region http://www.ncdc.noaa.gov/cag. For each year, we calculated the mean and spatial standard deviation of z-scores across available sites.

2.7 Separating spatial and temporal relationships

We separated spatial patterns across sites from site-level temporal variability using linear fits between a driver and response variable (e.g., precipitation and productivity). This spatial-temporal separation approach differs from the common practice in synthesis studies of eddy covariance data in which a single relationship between two variables is determined across pooled site years (Baldocchi, 2008; Law et al., 2002; Luysaert et al., 2007; Yu et al., 2013), representing the approximate spatial relationship. One interpretation of the separated temporal and spatial patterns is that site-level temporal slopes represent ecophysiological responses to annually varying factors such as precipitation, while spatial relationships fit to mean annual values across sites reflect slow-changing controls such as plant community adaptation to long-term climate (Biederman et al., 2016; Chen et al., 2015; Lauenroth & Sala, 1992).
RESULTS

3.1 | Mediterranean

The six sites in southern California (Figure 1, Table 1) have a Mediterranean climate with hot, dry summers (June–October) and cool, wet winters (November–May) during which >65% of annual precipitation typically falls. The Mediterranean Mid-elevation subregion includes four sites in Köppen class Csa (temperate with dry, hot summers) comprising one oak-pine forest and three chamaephytic ecosystems with elevation ~1500–1700 m, MAP ~550–700 mm, and MAT ~13°C. Winter precipitation (November–February) stored as soil moisture or occasionally as snowmelt drives ET and GEP peaking in May (Figure 2a). Productivity is sometimes sustained through summer months, depending upon stored soil moisture from winter and variable summer rainfall.

Two lower-elevation (~1300 m) Mediterranean sites in class Bwh (arid, desert, hot summers) including pinyon-juniper and chamaephytic ecosystems are warmer (~14°C) and receive less precipitation (~400 mm). Peak ET and GEP occur earlier in the spring (March–April) than at the Mid-elevation sites and are followed by an early summer dry-down. Summer precipitation drives a second ET peak sometimes accompanied by a smaller GEP peak (Figure 2b), suggesting that a greater fraction of ET is abiotic evaporation during the hot summer months than during spring. In these Mediterranean subregions (Figure 2a, b), Both GEP and EVI show single peaks at mid-elevation and two peaks at low elevation, although the EVI signal amplitude is small, as shown by Sims et al. (2006).

3.2 | Colorado Plateau

This site is dry (MAP=250 mm) with precipitation split evenly between cold winters (~0–3°C) and hot summers (~25–30°C). Stored winter moisture drives spring peaks in ET and GEP (May), which are weakly captured in EVI (Figure 2c). Summer precipitation drives ET without accompanying GEP, suggesting most ET in summer is evaporation, not transpiration. This site was separated from other Bsk sites (dry steppe, cold winters) in Monsoon Low (Figure 2f) due to higher winter precipitation (~50% of annual as compared to ~20–30% in Monsoon Low) and associated predominance of spring productivity.

3.3 | Monsoon

The Monsoon High-elevation subregion (Figure 2d) is represented by a mixed conifer forest at 3000 m elevation with MAT ~3°C. Mean annual precipitation is ~730 mm, split equally between winter and summer, with 20–50% falling as snow. Spring snowmelt typically provides sufficient soil moisture recharge to sustain ET and productivity through May and June until the monsoon onset in July. EVI captures these single peaks in GEP and ET.

The Monsoon Mid-elevation subregion (Figure 2e) comprises ponderosa pine, pinyon pine, and juniper woodlands at 1900–2500 m. MAT ranges from 6 to 11°C and MAP varies from 350 to 600 mm, with 50–70% during summer and some winter snowfall. Stored winter moisture drives spring ET and GEP (March–May). Productivity declines with soil moisture during early summer and then resumes with arrival of monsoon precipitation in July–August. It is this double peak in mean seasonal fluxes which characterizes this subregion, despite disparate Köppen classes (Table 1). EVI peaks once, in July, and fails to capture the spring GEP/ET peak.

The Monsoon Low subregion comprises desert grassland, shrubland, savanna, and deciduous forest at 500 to 1600 m. Mean climate varies widely (MAT ~12–23°C, MAP ~250–720 mm). These sites are unified by the dominant role of summer precipitation (~65–85% of MAP) driving single peaks of ET, GEP, and EVI (Figure 2f). Classes include Bsk and Bsh (dry steppe and classified as k-cold or h-hot, although MAT does not affect seasonal flux timing).

3.4 | Baja California Sur

This sea-level subtropical desert scrub site (MX-lpa) represents the only long-term eddy flux measurements (Figure 2g) in this hot, dry subregion (MAT = 24°C, MAP = 167 mm). Annual precipitation is dominated by moisture from Pacific tropical storms around September, which can drive ecosystem exchanges through the warm winter until as late as April. Therefore, we used a modified ecohydrologic year running June–May for annual flux integration.

3.5 | Patterns and drivers of regional CO2 exchange

Across the dataset of 150 site-years, annual NEP varied between -550 and +420 gC m-2, and mean annual NEP varied between -350 and +330 gC m-2 among sites (Figure 3, Table S1). While several ecosystems were persistently sinks or sources of CO2, about half the sites pivoted between sink and source functioning. Monsoon High and Mid sites were persistent sinks, most Monsoon Low sites pivoted, and Mediterranean sites exhibited the full range of sink/source/pivot function. All IGBP classes showed instances of both sink and source behavior except for mixed forest and evergreen needleleaf forests, which were sinks (Fig. S3). NEP interannual variation showed no identifiable spatial pattern, with standard deviations ranging between 20 and 230 gC m-2 (Table S1). NEP interannual variability was highest at the Mediterranean sites (Table S2), which corresponded closely with mixed forest and closed shrubland classes (Table S3).

We evaluated the relationship of CO2 exchange with climatic variables using the gross fluxes GEP and RECO. Despite distinct seasonal dynamics across subregions (Figure 2), previous work has shown that annual climate values are strongly associated with annual CO2 exchange across ecohydrologic years (Biederman et al., 2016; Pereira et al., 2007; Scott et al., 2015; Thomas et al., 2009). Mean annual GEP showed negative spatial relationships (ANOVA F-test versus a constant model) to MAT (Figure 4, R2 = 0.28, p < .05) and mean annual vapor pressure deficit (VPD, R2 = 0.26, p < .05, not shown) and a positive relationship to MAP (R2 = 0.41, p < .01). To reduce possible effects of a negative relationship between MAP and
MAT (e.g., high-elevation sites are wetter and colder, Fig. S41), we repeated the analysis with ten lower-elevation sites restricted to 270 mm < MAP < 450 mm and again found negative temperature effects on GEP (Fig. S5). Mean annual ET described the most spatial variability in GEP ($R^2 = 0.62$, $p < .01$), as previously shown for subsets of sites in this region (Biederman et al., 2016; Scott et al., 2015). Mean annual ecosystem respiration (Reco) showed no spatial relationship with MAT, VPD, or MAP, but there was a positive spatial relationship between Reco and mean annual ET ($R^2 = 0.26$, $p < .01$), possibly due to increased substrate production associated with GEP.

Interannual relationships of gross CO2 fluxes with $T$, $P$, and ET were analyzed using the full dataset of annual deviations (denoted by $\Delta$) from site-specific multiannual means (Figure S5, all fits $p < 0.01$). We found negative temporal responses of $\Delta$GEP to $\Delta T$ (-139 g C m$^{-2}$ °C$^{-1}$, $R^2 = 0.23$) and $\Delta$VPD (-884 g C m$^{-2}$ kPa$^{-1}$, $R^2 = 0.34$, not shown) and a positive response to $\Delta P$ (0.51 g C m$^{-2}$ mm$^{-1}$, $R^2 = 0.28$). Annual $\Delta$ET explained more $\Delta$GEP variability (1.19 g C m$^{-2}$ mm$^{-1}$, $R^2 = 0.60$) than was explained by $\Delta P$. Annual $\Delta$Reco showed positive temporal slopes to $\Delta P$ (0.27 g C m$^{-2}$ mm$^{-1}$, $R^2 = 0.17$) and $\Delta$ET (0.58 g C m$^{-2}$ mm$^{-1}$, $R^2 = 0.33$). Annual $\Delta$Reco showed a weak negative relationship with $\Delta T$ (Figure S5d), meaning that warmer years had less respiration (-55 g Cm$^{-2}$ °C$^{-1}$, $R^2 = 0.09$). To reduce possible effects of a negative relationship between $\Delta T$ and $\Delta P$ (Fig. S4b, e.g., cooler years are wetter), we repeated the analysis restricted to years with $|\Delta P| < 50$ mm and again found negative temperature effects on both $\Delta$GEP and $\Delta$Reco (Fig. S6).

Relative interannual variation of water availability ($P$, ET) and CO2 fluxes (GEP, Reco) was largest at sites with lowest water availability (mean ET) (Figure 6). Coefficient of variation of annual precipitation ($CV_P$) decreased from ~40% to 20% (not significant, $p = .14$, ANOVA $F$-test). $CV_{ET}$ was lower than $CV_P$ across the gradient with a declining trend from ~30% to 10% ($p < .01$). Of the measured fluxes, $CV_{GEP}$ decreased most steeply across the gradient, from ~60% at the driest sites to ~10% at the wettest sites ($p < .01$), while $CV_{Reco}$ showed a marginally declining trend from ~20% to ~10% ($p = .06$). Trends were computed excluding the Sky Oaks sites (US-sox), where 2005 precipitation was > two standard deviations above normal.

### 3.6 Regional variations of CO2 exchange across the Southwest: 2005–2014

The Southwest exhibited regionally coherent anomalies in annual variability of CO2 exchange (Figure 7). During 2005–2014, the portion of the study with greatest spatial representation of sites (Table 1), temporal precipitation anomalies ($zP$) were related to those of water availability ($zET$), productivity ($zGEP$), and net CO2 uptake ($zNEP$) (Table 2, Figure 7). The annual fluctuation of $zReco$ (not shown) was similar to $zGEP$.

The El Nino year 2005 (http://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/) was extraordinarily wet at the Mediterranean sites,
FIGURE 4  Spatial relationship of long-term mean annual GEP (top row a–c) and \( R_{\text{eco}} \) (bottom row d–f) with mean annual temperature (a, d), mean annual precipitation (b, e), and mean annual evapotranspiration (c, f). All slopes shown are significantly different from zero (panel a: \( p < .05 \); panels b, c, f: \( p < .01 \), \( n = 25 \) sites).

FIGURE 5  Temporal variations expressed as annual deviations from each site’s long-term mean values of the same variables shown in Figure 4. All linear fits shown are significant (\( p < .01 \)). \( n = 150 \) site-years.
FIGURE 6 Coefficients of interannual variation (CV) of (a) \( P \) and ET, and (b) GEP and \( R_{\text{eco}} \) across the gradient of site water availability (mean ET). Solid lines represent significant trends \( (p < .01) \). Dotted lines are not significant \( (p = .18 \) for precipitation, \( p = .13 \) for \( R_{\text{eco}} \)). Open symbols are for the Sky Oaks cluster (sob, son, soy), where 2005 precipitation was >2 standard deviations above normal.

FIGURE 7 Normalized annual anomalies (z-scores) of (a) precipitation (P), (b) evapotranspiration (ET), (c) gross ecosystem production (GEP), and (d) net ecosystem production (NEP). Shown are the mean and spatial standard deviation of z-scores across sites. The 10-year time period shown represents years with simultaneous measurements available for at least six sites representing at least four subregions (Table 1).
with some precipitation totals > 300% of normal (Figure 7a). However, the wet 2005 precipitation anomaly was weaker elsewhere in the Southwest, as reflected in the large spatial standard deviation of zP. The regional mean GEP and NEP in 2005 were close to average (mean $z_{\text{GEP}} = 0$, mean $z_{\text{NEP}} = -0.1$, Figure 7b–d). The El Niño year of 2010 was associated with increased precipitation across the Southwest, with $z_{\text{P}} = +0.9$ (spatial SD = 0.5) driving positive anomalies for ET, GEP, and NEP ($z_{\text{NEP}} = 1.1$ (0.4). In contrast to the regionally coherent precipitation increase of 2010, the 2011 La Niña year had spatially variable precipitation anomalies, reflected by a large standard deviation of precipitation $z_{\text{P}} = -0.4$ (1.1) (Figure 7a). 2011 Mediterranean site precipitation was above average ($z_{\text{P}} = +0.5$), while Monsoon site precipitation was below average ($z_{\text{P}} = -0.5$), resulting in regional mean ET, GEP, and NEP that were all close to normal (Figure 7b–d).

### 3.7 Comparing measured annual CO₂ fluxes and ET with MODIS-based models

MODIS GEP matched the spatial variability in measured mean annual GEP (Figure 8a), with a spatial slope of 1.04 and average bias of +81 gCm⁻² ($R^2 = 0.67$, $p < .05$). However, the slope between interannual deviations of MODIS GEP and tower-measured GEP across the full dataset was only 0.31 ($R^2 = 0.37$, $p < .05$), meaning that MODIS GEP captured less than one-third of the magnitude of interannual variability (Figure 8a inset), MODIS ET consistently underestimated long-term mean ET by an average of 50% (Figure 8b, spatial relationship: MODIS ET = 0.51 × Tower ET + 23; $R^2 = 0.52$, $p < .05$). Of note, mean measured and modeled ET matched better at one outlier site (MX-tes) for currently unknown reasons. MODIS ET captured only 29% of the magnitude of interannual variability in the tower-measured ET (Figure 8b inset, $R^2 = 0.45$, $p < .05$).

### 4 DISCUSSION

The extensive Southwest regional flux dataset presented here reveals unique characteristics of dryland CO₂ exchange and ET relative to wetter regions. Despite unique seasonal dynamics across climatic subregions (Figure 2), annual water availability was sufficient to predict much of the spatial and temporal variation of gross productivity (~60%) and respiration (~30–40%) (Figures 4 and 5). Greater temperature had negative effects on CO₂ exchanges (Figures 4 and 5), in contrast to positive effects in wetter systems. The ranges of mean annual CO₂ exchange and ET (Figures 3 and 4) and their interannual variability (Figures 5–7) were greater than for mesic regions, supporting inferences about the important role of semiarid regions in the global carbon cycle. However, MODIS-based models underestimated the interannual variability of GEP and ET and underestimated multiyear mean ET by 50% (Figure 8), suggesting that measuring and modeling interannual variation of dryland ecosystems...
Seasonal dynamics of subregions across the Southwest

Differences in seasonal timing of water availability and plant phenology (Figure 2) imply that dryland ecosystems in each subregion may respond differently to specific characteristics of climate change, such as seasonal precipitation shifts (Wolf et al., 2016). Apparent lags of up to several months between precipitation and ecosystem ET and CO₂ exchanges (Figure 2) suggest seasonal water storage was most important at Mediterranean sites and high-elevation sites, where winter precipitation is asynchronous with the warm temperatures and phenology conducive to ecosystem exchanges (Figure 2a, b, d, e) (Scott et al., 2012; Villarreal et al., 2016). An important implication is that seasonal-to-annual-scale analyses should be integrated over ecohydrologic years that pair incoming precipitation with resulting ecosystem exchange rather than using calendar-year values. This has been shown previously (Ma et al., 2007; Thomas et al., 2009) but not adopted as common practice in annual flux integrations. EVI coupling with ecosystem exchanges appeared to be weaker at sites where winter moisture is an important driver, such as the spring peaks in juniper, pinyon pine, and ponderosa pine ecosystems (Figure 2e), possibly because spring exchanges are dominated by evergreen overstory vegetation not showing strong changes in greenness (Barnes et al., 2016; Walther et al., 2016).

Spatial variability: long-term mean ET and carbon sink/source function

The wide range of CO₂ source and sink functioning (mean annual NEP of -350 to +330 gCm⁻² Figure 3, Table S1) contrasts with better-studied mesic to humid ecosystems, which are usually sinks (Baldocchi, 2008; Chen et al., 2015; Law et al., 2002; Luyssaert et al., 2007). The range of sink/source functioning observed across the Southwest also contrasts with concerns that eddy covariance data are suspect in drylands based on a limited number of sites that showed only sink behavior (Schlesinger, 2017). Wide variability in multiannual mean NEP could reflect legacies of ecosystem disturbance (e.g., drought, fire, insect infestation, harvest, grazing), which alter slow-changing controls on CO₂ exchange such as plant community structure and soil biogeochemical pools (Amiro et al., 2010; Baldocchi, 2008; Biederman et al., 2016; Dore et al., 2012).

The spatial patterns found here between long-term mean fluxes and water availability were broadly similar to those reported in prior flux data syntheses at continental to global scales (Figure 4, Table S4). As in prior studies, we did not force these relationships through the origin. We suggest that a negative GEP intercept in the GEP vs. ET spatial relationship represents a threshold of annual water availability (ET ~ 80 mm, Figure 4c) below which no productivity occurs (Biederman et al., 2016; Noy-Meir, 1973). Although this intercept represents an extrapolation from the data, the inclusion of arid sites in this study allows such extrapolation to remain small. A linear relationship between GEP and ET implies constant marginal water use efficiency for annually available water increments above this threshold. Further work is needed to characterize the physical and biological controls on the value of this ET threshold for productivity.

Notably, GEP showed a negative spatial relationship with MAT, in contrast to positive relationships from syntheses predominated by wetter sites. We found no spatial relationship between R_{eco} and MAT, in contrast to positive relationships in two prior studies (Table S4). Our results imply that future warming and drying could work in parallel to reduce the land carbon sink in drylands (Anderson-Teixeira et al., 2011) or even change it to a source (Figure 3).

Interannual variability of dryland ecosystem exchange

Measured CO₂ exchanges suggest the Southwest is a hot spot for interannual variability not identified in previous global modeling or empirical data upscaling studies, which lacked sufficient dryland flux data (Jung et al., 2011; Poulter et al., 2014). We found interannual variability of GEP and R_{eco}, as high as 60% and 30%, respectively (Figure 6), with the highest variability at Mediterranean sites (Table S2) and in shrublands (Table S3), similar to the driest sites reported for China (Yu et al., 2013; see Table S1). Meanwhile, mesic forested sites in North America showed ~7% interannual variability (Keenan et al., 2012), while 39 sites weighted toward mesic ecosystems had interannual variability of 5 to 25% (Yuan et al., 2009), comparable to the evergreen needleleaf and mixed forest sites in this study (Table S3). These differences likely reflect that dryland water availability is inherently more variable over time than the limiting resources in mesic ecosystems, such as temperature or nutrients. More than half of Southwest sites in the present study pivoted between functioning as CO₂ sources in dry years to sinks in wet years, as previously suggested for isolated sites or small site clusters (Ma et al., 2007; Pereira et al., 2007; Scott et al., 2015). While prior mesic flux studies have shown that drought reduces GEP more strongly than R_{eco}, reducing the carbon sink (Schwalm et al., 2010), we show here a prevalence of dryland sites acting as carbon sources (Figure 3) in warm, dry years (Figure 5, Table 2), supporting the idea that drylands strongly influence temporal variations in global CO₂ (Ahlström et al., 2015).

Interannual variation in water availability drove spatially coherent Southwest regional anomalies in GEP and NEP (Figure 7). The correlation between annual anomalies in ET and productivity (gross and net) was stronger than between P and productivity (Table 2), demonstrating the value of ET as a metric of ecosystem-available water following variable hydrologic losses (Equation 1). The strong correlation between zGEP and zNEP implies that interannual variability in net CO₂ exchange (not necessarily NEP magnitude) can be predicted from gross uptake, due to the interannual coupling of GEP and R_{eco} found in both observations (Baldocchi, Sturtevant, & Contributors, 2015; Biederman et al., 2016; Waring, Landsberg, & Williams, 1998)
and models (Jung et al., 2011). In recent work, we showed that a change of ±100 mm in annual water availability results in an average change of ±6.4 g C m⁻² of NEP across a subset of semiarid flux sites (Biederman et al., 2016). NEP magnitude, however, also depends on slow-changing controls (i.e., soil biogeochemistry, plant community) (Biederman et al., 2016; Ma et al., 2016).

The negative interannual relationships of temperature with GEP and R_{eco} (Figure 5, Fig. S6) contrast with positive relationships expected from prior analyses that aggregated site-years (Luyssaert et al., 2007; Reichstein et al., 2007). To our knowledge, the interannual sensitivity of R_{eco} to temperature across a regional flux dataset has not previously been reported, although Yu et al. (2013) reported a positive spatial relationship (Table S4). In the water-limited Southwest, higher temperatures could increase abiotic evaporative losses at the expense of water for photosynthesis (Mekonnen, Grant, & Schwalm, 2016; Noy-Meir, 1973) and raise vapor pressure deficit (VPD), reducing plant water use efficiency (Beer et al., 2009; Zhou, Yu, Huang, & Wang, 2015). Meanwhile, direct temperature limitation of R_{eco} contrasts with conventional understanding of plant and microbial kinetics (Reichstein et al., 2005; Ryan & Law, 2005). We think it more likely that the negative temperature-R_{eco} relationships reflect a causal chain in which higher VPD reduces stomatal opening and further production of the photosynthesates which supply an array of autotrophic and heterotrophic respiration processes (Chen et al., 2015; Ma et al., 2007; Ryan & Law, 2005; Sampson, Janssens, Curiel Yuste, & Ceulemans, 2007; Vargas et al., 2011; Waring et al., 1998).

4.4 Comparison of remote sensing models with measured carbon and water exchanges

Our results suggest that RSM reasonably predict spatial patterns of average annual GEP across the Southwest (Figure 8a), similar to prior results in water-limited regions (Jin & Goulden, 2014; Verma et al., 2014). However, RSM systematically underestimated ET by 50% (Figure 8b), in contrast with Mu et al. (2011), who found 86% accuracy compared to a North American set of 46 eddy covariance towers. This bias suggests that MODIS ET overestimates water use efficiency (GEP/ET) by ~100%, supporting the idea that ET-GEP relationships in observational data should be used to constrain model estimates of each flux (Mystakidis, Davin, Gruber, & Seneviratne, 2016). Notably, RSM captured only 20–30% of measured interannual variability magnitude for both GEP and ET (Figure 8 insets, temporal slopes ~0.2–0.3). Larger interannual variability in the measurements implies semiarid regions such as the Southwest likely regulate global CO₂ variation even more strongly than suggested by model-based studies (Ahlström et al., 2015; Poulter et al., 2014). Furthermore, semiarid regional CO₂ exchange may be even more sensitive to precipitation anomalies (e.g., wet periods or droughts), than previously thought (Haverd, Ahlström, Smith, & Canadell, 2017; Poulter et al., 2014).

Several factors could underlie these significant RSM-data mismatches in dryland ecosystems. Dryland ecosystems of the Southwest show low spectral sensitivity to water stress (Figure 2), limiting the information content of EVI, FPAR, and/or LAI inputs to RSM (Ha et al., 2015; Sims, Brzostek, Rahman, Dragoni, & Phillips, 2014). Furthermore, sparse weather observational data may introduce errors in gridded meteorological inputs to RSM, especially in the Southwest, where weather may vary strongly over short temporal and spatial scales (Goodrich et al., 2008; Noy-Meir, 1973; Zhao, Running, & Nemani, 2006). However, a RSM-data comparison of daily ET (Mu et al., 2011) for a subset of three Southwest sites that were also included in this study suggests that inadequate model structure and/or parameterization was more important than the choice of weather forcing between tower and gridded datasets (Table S5: first 3 rows). Meanwhile, for one forested SW site (US-fuf) and the full dataset of 46 North American towers in the study of Mu et al. (2011), with heavy representation by forest sites, RSM-data agreement was impacted to similar degrees by model changes and weather forcing. RSM-data agreements for drylands are likely disproportionately influenced by model representation and parameterization of water constraints. MODIS GEP and ET estimates are based on only atmospheric VPD constraints and thus do not consider potential soil moisture constraints (Krofcheck et al., 2015; Liu et al., 2015; Sims et al., 2014; Verma et al., 2014). Recent work has demonstrated soil moisture is more important than VPD for control of stomatal conductance at dryland flux sites (Novick et al., 2016), suggesting added soil moisture constraints could significantly improve RSM-data agreement. Additionally, emerging satellite indices that better capture changes in plant physiological function (e.g., water stress), such as solar-induced fluorescence (SIF), could also improve RSM-data agreement (Zhang et al., 2014).

4.5 Directions for future work to improve semiarid CO₂ and water exchange estimates

The dryland flux dataset assembled here can help address a significant paradox: Model analyses indicate the importance of drylands globally, while the models remain under-constrained by data in water-limited regions (Ahlström et al., 2015; Jung et al., 2011; Poulter et al., 2014). Model-derived maps of ecosystem exchanges are needed to place regions within a global context to understand the impacts of vegetation change and land management, and to predict terrestrial feedbacks to climate change (Friedlingstein et al., 2013).

We suggest several activities for improved understanding of ET and CO₂ exchanges in drylands: (i) Continuous, long-term dryland flux datasets present opportunities for testing and improving data inputs and model structure for RSM. Specifically, work is needed to evaluate how moisture stress is represented (i.e., VPD vs. soil moisture control of stomatal conductance). (ii) Dryland flux datasets could be scaled regionally using data-driven approaches (Jung et al., 2011) to constrain models more broadly than at the 25 sites presented here, and/or as an alternative method of quantifying regional CO₂ exchanges through time by translating normalized anomalies (Figure 7) into mass fluxes (e.g., Haverd et al., 2013); and (iii) Regional datasets (present study, Haverd et al., 2013; Ma et al., 2013; Yu et al., 2013) should be used to enhance the representation of
semiarid regions within global datasets such as La Thuile 2007 and FluxNet 2015 (http://fluxnet.fluxdata.org/data/) enabling model-data fusion exercises across a broader range of gradients in environmental drivers.

ACKNOWLEDGEMENTS

Funding for J.A.B. and the AmeriFlux core sites run by R.L.S. and M.E.L. was provided by the U.S. Department of Energy’s Office of Science. USDA is an equal opportunity employer. Research at site US-so and was supported by the University of Utah and the USGS. S.A.P. acknowledges National Science Foundation grant EAR–125501. Research at San Diego State University Sky Oaks Biological Field Station (US-soo/soob, son, soy) was supported by the NSF SER, SDSU and SDSU Field Stations. Research in La Paz, BCS, MX (MX-lpa) was supported by NSF International Program, CONACyT, SDSU, and CIBNOR. We thank B. Poulier and three anonymous reviewers for feedback on earlier drafts of the manuscript and R. Bryant for extensive data processing. The data and analysis codes for the figures and tables in this study are available upon request from the corresponding author.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

J.A.B., R.L.S., and M.L.G. conceived the study. J.A.B., G.P.C., and D.K. assembled the data; J.A.B. produced preliminary results and wrote the manuscript. All authors analyzed data, contributed to the interpretation of results, and helped revise the manuscript.

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.