

Robust inference of ecosystem soil water stress from eddy covariance data

Brandon P. Sloan ^{a,b,c,*}, Xue Feng ^{b,c}

^a Bioresearch Science and Engineering Group, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States

^b Department of Civil, Environmental, and Geo- Engineering, University of Minnesota - Twin Cities, Minneapolis, MN, United States

^c St. Anthony Falls Laboratory, University of Minnesota - Twin Cities, Minneapolis, MN, United States

ARTICLE INFO

Dataset link: <https://github.com/sloan091/afm-robust-stress-id>

Keywords:

Soil water stress
Drought
Plant water use strategy
Ecosystem stress
Eddy covariance
FLUXNET2015
AmeriFlux
Robustness

ABSTRACT

Eddy covariance data are invaluable for determining ecosystem water use strategies under soil water stress. However, existing stress inference methods require numerous subjective data processing and model specification assumptions whose effect on the inferred soil water stress signal is rarely quantified. These uncertainties may confound the stress inference and the generalization of ecosystem water use strategies across multiple sites and studies. In this research, we quantify the sensitivity of soil water stress signals inferred from eddy covariance data to the prevailing data and modeling assumptions (i.e., their robustness) to compile a comprehensive list of sites with robust soil water stress signals and assess the performance of current stress inference methods. To accomplish this, we identify the most prevalent assumptions from the literature and perform a digital factorial experiment to extract probability distributions of plausible soil water stress signals and model performance at 151 FLUXNET2015 and AmeriFlux-FLUXNET sites. We develop a new framework that summarizes these probability distributions to classify and rank the robustness of each site's soil water stress signal, which we display with a user-friendly heat map. We estimate that only 5%–36% of sites exhibit a robust soil water stress signal due to deficient model performance and poorly constrained ecosystem water use parameters. We also find that the lack of robustness is site-specific, which undermines grouping stress signals by broad ecosystem categories or comparing results across studies with differing assumptions. Lastly, existing stress inference methods appear better suited for eddy covariance sites with grass/annual vegetation. Our findings call for more careful and consistent inference of ecosystem water stress from eddy covariance data.

1. Introduction

Eddy covariance observations are invaluable for characterizing ecosystem water use strategies under soil water stress. The hundreds of eddy covariance sites maintained by observation networks across the globe (e.g., AmeriFlux, ICOS, AsiaFlux, TERN-OzFlux) have been leveraged by researchers to generalize ecosystem water use strategies across plant functional types (PFTs) and climatic gradients (Zhou et al., 2014, 2015; Lin et al., 2018; Boese et al., 2019; Fu et al., 2022a), identify soil moisture thresholds for ecosystem stress responses (Bassiouni et al., 2018; Fu et al., 2022a,c), and quantify the relative importance of soil versus atmospheric water stress (Novick et al., 2016; Kimm et al., 2020; Wang et al., 2022; Fu et al., 2022c; Liu et al., 2020). Yet, the growing body of eddy covariance research has not comprehensively answered a critical question: which eddy covariance sites have actually observed ecosystem responses to soil water stress?

In the context of eddy covariance data, soil water stress typically refers to reductions in evapotranspiration (ET) and/or gross primary

productivity (GPP) with respect to drying soil caused by the stomatal closure of the plants in the ecosystem. However, many eddy covariance sites have likely not encountered a significant period of soil water stress given their location in hydric or mesic ecosystems that may only experience intermittent drought, which are likely missed by the limited eddy covariance observation periods (e.g., Fig. 1 in Pastorello et al. (2020)). Therefore, it is difficult to know *a priori* if a site has experienced soil water stress without doing extensive literature review, and, even then, many sites do not have studies pertaining to soil water stress. Most studies on ecosystem responses to soil water stress that leverage multiple eddy covariance sites do not exclude non-stressed sites, which may confound their generalizations of ecosystem water use strategies. Consistent analysis across eddy covariance sites and ecosystems requires a comprehensive list of sites that have experienced soil water stress, which leads to a second unanswered question: how well can we infer soil water stress signals from eddy covariance data?

* Corresponding author at: Bioresearch Science and Engineering Group, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States.

E-mail address: sloanbp@ornl.gov (B.P. Sloan).

<https://doi.org/10.1016/j.agrformet.2023.109744>

Received 2 March 2023; Received in revised form 19 September 2023; Accepted 27 September 2023

Available online 4 November 2023

0168-1923/© 2023 Elsevier B.V. All rights reserved.

Table 1

Summary of the key data uncertainty and modeling specification assumptions found in studies inferring ecosystem water use strategy from eddy covariance data. Each column represents a specific assumption group or treatment discussed in Section 2.3.1, which are used to distill the assumption sets or levels shown in Table 2.

Study	1. Fit Algorithm	2. SEB ^a	3. Growing Season	4. LAI	5. G _c Equation	6. Fitting Parameters	7. Response Variable	8. VPD
Zhou et al. (2014)	Grid search	-	DOY and/or DD GPP > 30% of max GPP each year	-	uWUE ^c	m	-	VPD _a
Zhou et al. (2015)	Zhou et al. (2014)	-	HH GPP > 10% max GPP each year	-	uWUE ^c	m	-	VPD _a
Novick et al. (2016)	OLS	-	T _a – based DOY period	-	Oren et al. (1999)	G _{s,ref} , m	G _c	VPD _a
Medlyn et al. (2017)	NLLS	-	3 most productive months; 09:00-15:00	Max LAI	Eqn. 2	G ₁ , G ₀ = 0	G _c	VPD _a
Knauer et al. (2017)	Robust NLLS	H, LE, R _n - G, BR	15 day moving average > 50% of 95th DD GPP percentile	Max LAI	Eqn. 2	G ₁ , G ₀ = 0	G _c	VPD _a , VPD _l
Lin et al. (2018)	NLLS w/ bootstrap	-	GPP > 10% of 95th percentile HH GPP data.	PAR-derived using Xu et al. (2010)	Eqn. 1 ^d	G ₀ , G ₁ , m	G _c	VPD _l
Li et al. (2018)	NLLS	H, LE, R _n - G	Lin et al. (2018)	Same as Lin et al. (2018)	Eqn. 1	G ₀ , G ₁ , m	G _c	VPD _l
Nie et al. (2021) ^e	Li et al. (2019)	-	Knauer (2017)	-	Eqn. 1	G ₀ , G ₁ , m	-	VPD _l
Boese et al. (2019) ^e	NLLS w/ grid search	-	GPP > 0.1 gC/d/m ²	-	Modified Zhou et al. (2015)	-	ET (daily)	VPD _a
Kimm et al. (2020) ^e	OLS/NLLS	-	July–August	-	Multiple regression and Leuning (1995)	-	G _c (daily)	VPD _a
Bassiouni et al. (2021)	Constrained NLLS w/ bootstrap	-	GPP > 25% of 95th GPP percentile	Assumed GPP captures LAI	Eqn. 2	$\lambda \propto 1/G_1, G_0$	ET	VPD _a
Fu et al. (2022) ^{b,e}	NLLS	BR	DD Ta > 15C, DD VPD _a > 0.5 kPa, DD S _{w,in} > 250 W/m ²	Seasonal detrend	Used ANN and multiple regression	-	G _c (daily)	VPD _a

Abbreviations: Ordinary least squares (OLS), Nonlinear least squares (NLLS), Not addressed (-), Latent heat flux (LE), Net radiation (R_n), Ground heat flux (G), Bowen Ratio (BR), Gross primary productivity (GPP), Leaf Area Index (LAI), Day of year (DOY), Daily average (DD), Half-hourly (HH), Artificial Neural Network (ANN), Underlying water use efficiency (uWUE), Intercept parameter (G₀), Slope parameter (G₁), VPD exponent (m), Canopy conductance (G_c), Evapotranspiration (ET), atmospheric vapor pressure deficit (VPD_a), leaf-to-air VPD (VPD_l).

a. Entries refer to which flux the SEB errors were applied to (see Sect. S5).

b. Also checked the three most productive GPP months for growing season.

c. The uWUE model uses the same underlying theory of Eq. 2 of this paper.

d. The authors also tested the Oren et al. (1999), Leuning et al. (1995), and Lloyd et al. (1994) model.

e. PMOC model not explicitly used in this paper.

Most studies inferring ecosystem soil water stress signals from eddy covariance data follow a similar workflow: select relevant eddy covariance sites, fit an interpretable model to the data to infer the soil water stress signal, and group site-level signals by plant functional types (PFTs) and/or climatic indices to search for general patterns across ecosystem types. Within this workflow lies numerous subjective assumptions to control for the eddy covariance data uncertainty (e.g., noise, phenology) and specify an interpretable model (e.g., statistical, mechanistic). Unfortunately, these assumptions vary across studies (e.g., Table 1), with minimal quantification of their effects on the inferred soil water stress signal (see Knauer et al. (2018) for a counter-example). Therefore, the sensitivity of the inferred ecosystem soil water stress signal to the numerous data and modeling assumptions—i.e., its robustness—must also be quantified to confidently identify sites with soil water stress.

In summary, the ability to extract meaningful conclusions about ecosystem responses to soil water stress from eddy covariance data requires a more thorough understanding of the presence and robustness of the inferred soil water stress signals. The goals of this study are to: (1) develop a framework to quantify and rank the robustness of soil water stress signals inferred from eddy covariance data to common data and modeling assumptions, (2) provide a comprehensive list of

eddy covariance sites with robust soil water stress, (3) examine how non-robust soil water stress signals confound generalizations across ecosystem categories, and (4) identify deficiencies in current soil water stress inference approaches.

To achieve these goals, we have compiled a list of key data and modeling assumptions taken from studies of ecosystem water use strategies inferred from eddy covariance data. Then, for each of 151 FLUXNET2015 and AmeriFlux-FLUXNET sites, we construct a probability distribution of inferred soil water stress signals and model performance due to the range of plausible assumption sets. We explore these stress signal distributions in a case study of nearly identical eddy covariance sites with known soil water stress to illustrate the complexity of robust soil water stress inference (Section 3.1). Next, we assess how the robustness at all 151 sites affects detecting patterns in soil water stress signals grouped across broader ecosystem categories (Section 3.2). Finally, we propose a robust ecosystem soil water stress framework that summarizes the stress signal and model performance distributions to create a rank-ordered list of eddy covariance sites with robust soil water stress (Section 3.3).

This research is fitting for the 25th Anniversary of AmeriFlux special issue as it assesses the current practices of inferring ecosystem soil water stress signals from growing eddy covariance datasets. To

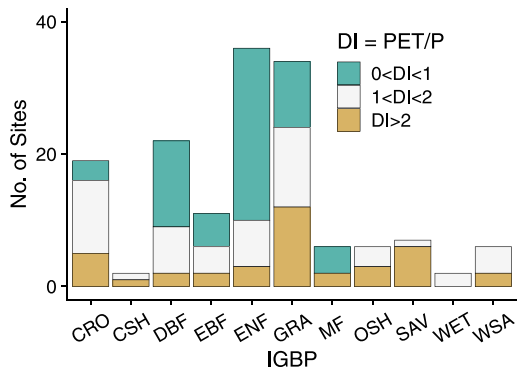


Fig. 1. 151 eddy covariance sites taken from the FLUXNET2015 and AmeriFlux-FLUXNET data sets (Pastorello et al., 2020) used in this analysis, categorized by the International Geosphere-Biosphere Programme (IGBP) plant functional types (PFTs) and the annual Dryness Index (DI = Potential ET/Precipitation).

our knowledge, this research is the first attempt at comprehensively identifying which eddy covariance sites exhibit robust soil water stress signals while highlighting deficiencies in current inference approaches. We hope these results will inspire further work by AmeriFlux and other eddy covariance communities to create guidance for robust soil water stress inference across studies and ecosystems.

2. Methods

2.1. Common soil water stress inference workflow

In the following subsections, we lay out our workflow for inferring soil water stress from eddy covariance data that we have summarized from the studies listed in Table 1.

2.1.1. Eddy covariance site selection and baseline filtering

To create a comprehensive list of eddy covariance sites with robust soil water stress signals, we first downloaded half-hourly/hourly eddy covariance data from both the FLUXNET2015 (Pastorello et al., 2020) and AmeriFlux-FLUXNET data products. We selected 151 of the 229 potential sites that had adequate soil moisture and atmospheric observations (e.g., net radiation) relevant to the interpretable model discussed in Section 2.1.2. The 151 eddy covariance sites cover a range of ecosystems (Fig. 1). For each site, we selected the soil moisture sensor with the best coverage (typically shallower) and collected ancillary data (e.g., vegetation height, tower height) from metadata or literature. The full site and sensor selection details are given in Section S1, and a summary table of the 151 eddy covariance sites is given in Table S1.

In this study, we aim to test the sensitivity of the inferred soil water stress signals to common data uncertainty assumptions that are applied inconsistently across studies (see Section 2.3). However, many of the studies in Table 1 agree on some baseline filtering assumptions that focus the analysis on periods where vegetation modulates carbon and water fluxes to the atmosphere. We apply these baseline filtering assumptions (listed in Fig. 2a and discussed in Sect. S1) to all 151 eddy covariance sites.

2.1.2. Interpretable model specification

The existing literature uses a wide range of interpretable models to infer soil water stress signals from eddy covariance data, ranging from statistical (Koster et al., 2009; Fu et al., 2022a) to mechanistic (Novick et al., 2016; Lin et al., 2018). Here, we adopt a popular mechanistic approach that combines the Penman–Monteith equation (Monteith, 1965) with an optimal canopy conductance (G_c) equation to create a simplified land surface model. These Penman–Monteith Optimal Conductance models (called PMOC from hereon) provide an advantage

over statistical models given their basis in first principles (e.g., conservation of mass and energy) and interpretable fitting parameters linked theoretically and empirically to ecosystem water use strategies.

The Penman–Monteith equation (Eq. (1)) balances the effect of the surface energy balance (SEB) and atmospheric turbulence on evapotranspiration (ET , $\text{mm} \cdot \text{d}^{-1}$) from a plant canopy, where Δ ($\text{Pa} \cdot \text{K}^{-1}$) is the slope of the Clausius–Clapeyron relationship, R_n ($\text{W} \cdot \text{m}^{-2}$) is the net radiation, G ($\text{W} \cdot \text{m}^{-2}$) is the ground heat flux, ρ_a ($\text{kg} \cdot \text{m}^{-3}$) is the air density, c_p ($\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$) is the specific heat of air at constant pressure, G_a ($\text{m} \cdot \text{s}^{-1}$) is the atmospheric conductance, VPD_a (Pa) is the atmospheric vapor pressure deficit, γ ($\text{Pa} \cdot \text{K}^{-1}$) is the psychrometric constant, and c_1 (≈ 0.035) converts LE ($\text{W} \cdot \text{m}^{-2}$) to ET . Except for G_c , the inputs to Eq. (1) are constants or taken from eddy covariance observations (see Sect. S2 for full details).

$$ET = c_1 \cdot \frac{\Delta \cdot (R_n - G) + \rho_a \cdot c_p \cdot G_a \cdot VPD_a}{\Delta + \gamma \cdot \left(1 + \frac{G_a}{G_c}\right)} \quad (1)$$

The canopy conductance (G_c , $\text{m} \cdot \text{s}^{-1}$) represents how the ecosystem modulates ET through stomatal control on transpiration and the decline in surface moisture on evaporation. We require an additional equation for G_c that partitions evaporation and transpiration from the canopy as well as the response of stomata to other environmental variables (e.g., VPD). Classic plant-scale optimality theory provides numerous formulations for stomatal conductance (g_s) based on the assumption that plants close their stomata to maximize carbon gain for a given amount of water loss (Cowan and Farquhar, 1977). These optimal g_s equations are then scaled to the optimal G_c equations using ecosystem scale observations. For example, the following G_c formulation from Lin et al. (2018) is a scaled version of the g_s formulation from Lloyd et al. (1994):

$$G_c = \left[G_o + G_1 \cdot \frac{GPP}{C_a \cdot VPD^m} \right] \cdot c_2 \quad (2)$$

where GPP ($\mu\text{moles CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$), atmospheric CO_2 concentration (C_a , $\mu\text{moles CO}_2 \cdot \text{moles air}^{-1}$), and VPD (kPa ; can differ from VPD_a ; see Sect. 2.3.1) are eddy covariance observations, and G_o ($\mu\text{moles air} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$), G_1 (kPa^m), and m are fitting parameters. The term c_2 (≈ 0.025) converts G_c from a molar flux to a volume flux (Lin et al., 2018). The intercept parameter, G_o , represents the soil/plant surface evaporation (Li et al., 2019), and the exponent parameter, m , describes the sensitivity of stomatal closure to VPD (across plants in the canopy) with an optimal value of 0.5. The slope parameter, G_1 , is theoretically and empirically linked to ecosystem water use strategy through its inverse relationship to the marginal water use efficiency (λ) as higher (lower) G_1 indicates plants are more aggressive (conservative) in keeping their stomata open to assimilate CO_2 at the cost of transpiration (Cowan and Farquhar, 1977; Lloyd et al., 1994; Medlyn et al., 2011). Therefore, G_1 is the critical parameter for inferring ecosystem responses to soil water stress. In this study, we use an additional G_c formulation from Medlyn et al. (2017) shown in Equation S5.

2.1.3. Soil water stress inference

We use the relationship between the ecosystem water use strategy parameter (G_1) and soil moisture to represent the ecosystem soil water stress signal based on PMOC studies (Table 1) and decades of theory and experiments (Cowan and Farquhar, 1977; Hari et al., 1986; Lloyd et al., 1994; Manzoni et al., 2011; Zhou et al., 2013; Wolf et al., 2016; Drake et al., 2017). A positive (negative) relation between G_1 and soil moisture (red (blue) solid line in Fig. 2f) indicates dry (wet) soil water stress, where plants appear to close stomata in response to lower leaf water potential (lack of oxygen) caused by drying (water-logged) soil. For each eddy covariance site (and unique assumption set), we fit the PMOC model to the data split into 10 soil moisture percentiles bins (θ_p), resulting in at most 10 estimates of G_1 (e.g., red and blue dots in Fig. 2f). Then, we fit a segmented or straight line regression model to

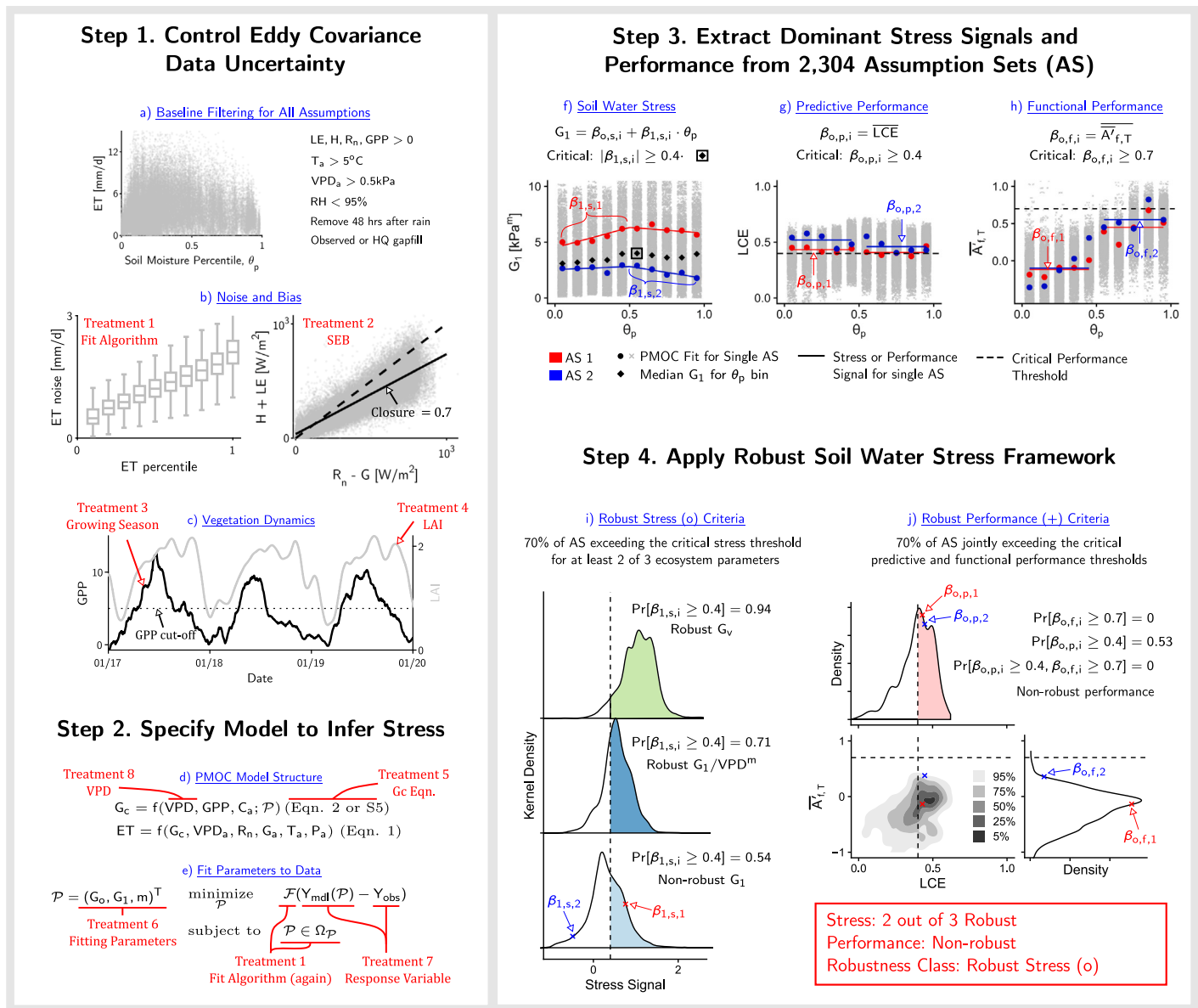


Fig. 2. Workflow for quantifying and ranking the robustness of the ecosystem soil water stress signal inferred from each eddy covariance site, using the (a) US-Me2 eddy covariance site as an example. Step 1 illustrates Treatments 1–4 in Table 2 that control data uncertainty due to (b) noise and surface energy budget (SEB) closure bias and (c) vegetation dynamics that are handled with a “growing season” filter (i.e., keep days above GPP cut-off) or by augmenting the data with MODIS LAI. Step 2 locates where Treatments 5–8 in Table 2 occur in the (d) Penman–Monteith Optimal Conductance (PMOC) model and (e) parameter estimation. See Sect. S5 for full details on the treatment levels. Step 3 illustrates the (f) soil water stress signal extraction for two of the 2,304 assumption sets (AS, gray x’s), which differ in their Treatment 5–6 assumptions. The location of the extracted stress signals is used to extract both the mean (g) predictive and (h) functional performance for each AS. Finally, Step 4 illustrates how probability distributions of the 2,304 (i) soil water stress signals and (j) performance metrics are summarized based on a satisficing metric (dashed black line) to classify each site’s robustness (see Section 2.3.2).

the $\theta_p - G_1$ points and extracted the dominant ecosystem soil water stress signal (slope $\beta_{1,s,i}$ in Fig. 2f). The stress signal detection was performed in R version 4.1.0 (R Core Team, 2021) using the *segmented* package (Mugge, 2008). See Section S3 for full details on the dominant stress signal extraction.

In addition to G_1 , we also check two other ecosystem parameters—the $\frac{G_1}{VPD^m}$ and the vegetation conductance $G_v (= \frac{G_1 \cdot GPP}{VPD^m})$ —to improve the chances of detecting a soil water stress signal. The parameter G_1 may be confounded by parameter correlations with G_o and m during the estimation process (i.e., practical identifiability issues (Guillaume et al., 2019)); therefore, both $\frac{G_1}{VPD^m}$ and G_v can counter correlations between G_1 and m , while G_v may incorporate non-stomatal limitations (Zhou et al., 2013; Dewar et al., 2018) and drought-deciduousness (Novick et al., 2019) through its inclusion of GPP. However, these two parameters have the disadvantage of potentially misattributing seasonality in

VPD (e.g., land–atmosphere interactions (Seneviratne et al., 2010)) and GPP (e.g., phenology) as soil water stress. Therefore, we check the agreement between soil water stress signals derived from G_1 , $\frac{G_1}{VPD^m}$, and G_v in the later analysis of robustness of soil water stress at each site (see Section 2.3.2). We use the median VPD and GPP in each θ_p bin along with the estimated G_1 to calculate $\frac{G_1}{VPD^m}$, and G_v .

2.1.4. Ecosystem generalization

Many PMOC studies (Table 1) tend to compare ecosystem water use strategies by grouping soil water stress signals (or other inferred ecosystem water use parameters) from multiple eddy covariance sites by dominant vegetation type and/or climatic indices. Here, we use the International Geosphere–Biosphere Programme (IGBP) plant functional types (PFTs) and the annual Dryness Index (DI = Potential Evapotranspiration/Precipitation) to define our ecosystem categories (Fig. 1). The

Table 2

The data uncertainty and modeling specification treatments (columns) and levels (rows) used to create the probability distribution functions of the soil water stress signals and performance metrics at each eddy covariance site (e.g., Fig. 2i–j). These levels were derived from the literature review in Table 1.

	1. Fit Algorithm ^a	2. SEB	3. Growing Season	4. LAI	5. G _c Eqn.	6. Fitting Parameters	7. Response	8. VPD
1	NLLS	None; assumes all SEB error is from H	None	None	Lin et al. (2018); Eqn. 1	G ₀ , G ₁ , m	G _c	VPD _a
2	Robust NLLS (iterative reweighting)	Use BR corrected LE standard in FLUXNET products ^b	Remove days where DD GPP < 50% of the 95 th percentile of 15-day smoothed DD GPP ^c	Normalize fluxes by smoothed MODIS LAI ^d	Medlyn et al. (2017); Eqn. 2	G ₁ , m	ET	VPD _i calculated from Eqn. 3
3	Constrained L1-weighted SSE optimization	Add SEB residual to LE	-	-	-	G ₀ , G ₁	-	Estimate VPD _i and response variable (G _c or ET)
4	Constrained LCE optimization	-	-	-	-	G ₁	-	-

Abbreviations: Nonlinear least squares (NLLS); Sum of squared errors (SSE), Lee and Choi Efficiency (LCE), Surface Energy Budget (SEB), Latent heat flux (LE), Bowen Ratio (BR), Gross primary productivity (GPP), Leaf Area Index (LAI), Daily average (DD), Not addressed (-), Intercept parameter (G₀), Slope parameter (G₁), VPD exponent (m), Canopy conductance (G_c), Evapotranspiration (ET), atmospheric vapor pressure deficit (VPD_a), leaf-to-air VPD (VPD_i)

a. The optimization methods differ from least squares in that the objective value is a scalar (SSE or LCE) rather than a vector of residuals. For least squares methods we used *fitlm* and for optimization we used *fmincon* in MATLAB.

b. The FLUXNET method does not force closure at each time step, but rather calculates an average correction factor over a multi-day window (Pastorello et al., 2020).

c. Same method as Knauer et al. (2017).

d. See Sect. S6 for details.

IGBP PFTs are Evergreen Needleleaf Forest (ENF), Mixed Forest (MF), Deciduous Broadleaf Forest (DBF), Evergreen Broadleaf Forest (EBF), Woody Savanna (WSA), Grassland (GRA), Closed Shrubland (CSH), Cropland (CRO), Savanna (SAV), Open Shrubland (OSH), and Wetland (WET). We calculated the annual DI using the CRU TS v4.03 (Harris et al., 2020) rather than the eddy covariance data given their short observation periods. Next, we discretized the DI into three dryness regimes: hydric (DI < 1), mesic (1 < DI < 2), and xeric (DI > 2). In this study, the term “ecosystem” describes the spatial scale of eddy covariance observations as well as a broad category defined by PFT and DI. We test the efficacy of these broad ecosystem categories in detecting general soil water stress patterns in Section 3.2.

2.2. Predictive and functional performance

We quantify the predictive and functional performance of the PMOC model to check the quality of the soil water stress inference (e.g., a strong stress signal with weak performance should be interpreted with caution). Here, the predictive performance refers to the ability of the PMOC model to match the data, whereas the functional performance refers to whether the model matches the data for the right reasons.

For predictive performance, we use the Lee and Choi Efficiency (LCE; Eq. (3) (Lee and Choi, 2022)), which is a re-balanced formulation of the Kling–Gupta Efficiency (Gupta et al., 2012) that better captures the model to observation correlation coefficient (r), standard deviation ratio (α), and mean ratio (β_μ). A value of 1 indicates a perfect match and there is no lower bound.

$$LCE = 1 - \sqrt{(r \cdot \alpha - 1)^2 + (r/\alpha - 1)^2 + (\beta_\mu - 1)^2} \quad (3)$$

For functional performance, we use information theory to quantify the ability of a model to match the strength of relationship between predictor variable(s) and the response variable (Ruddell et al., 2019; Bassiouni and Vico, 2021). Here, our functional performance metric ($A'_{f,T,i}$; Eq. (4)) is a re-scaled relative difference in the multivariate total mutual information between model and observations proposed by Ruddell et al. (2019). The total mutual information ($I(X_1, X_2; Y)$; Eq. (5)) estimates the information about a response variable (Y) contained in predictor variables (X) using Shannon's Entropy (H). Therefore, $A'_{f,T,i}$ measures how closely the model matches the information transfer between chosen predictor and response variables in the observations.

We select three predictor variables (VPD_a , θ_p , and GPP) and calculate the metric for each pair in each θ_p bin. The values of $A'_{f,T,1}$, $A'_{f,T,2}$, and $A'_{f,T,3}$ correspond to $A'_{f,T}(\theta_p, VPD_a; Y)$, $A'_{f,T}(\theta_p, GPP; Y)$, and $A'_{f,T}(GPP, VPD_a; Y)$, where Y represents either ET or G_c based on the response variable selection (see Treatment 7 in Section 2.3.1). See Section S4 for calculation details for $A'_{f,T,i}$. We use the individual $A'_{f,T,i}$ values to assess site specific results in this paper, but the average of the three functional metrics ($\bar{A}'_{f,T}$) is used to classify robustness (see Section 2.3.2).

$$A'_{f,T}(X_1, X_2; Y) = 1 - |A_{f,T}(X_1, X_2; Y)| \\ = 1 - \left| \frac{I(X_1, X_2; Y_o) - I(X_1, X_2; Y_m)}{I(X_1, X_2; Y_o)} \right| \quad (4)$$

$$I(X_1, X_2; Y) = H(X_1, X_2) + H(Y) - H(X_1, X_2, Y) \quad (5)$$

2.3. Quantifying robustness of ecosystem soil water stress signals

In the following subsections, we explain how we quantify the robustness of soil water stress signals to common, subjective data and modeling assumptions (Section 2.3.1) and lay out our robustness framework for ranking the soil water stress signals of the 151 eddy covariance sites (Section 2.3.2).

2.3.1. Data uncertainty and model specification assumptions

We reviewed the growing literature using PMOC and similar approaches to infer ecosystem water use strategies from eddy covariance data to find common assumption groups or “treatments” (columns of Table 1). We selected treatments that were applied inconsistently across studies (rows of Table 1) without definitive guidance on what constitutes best practice. We selected 2–4 assumptions or “levels” for each treatment (Table 2) to test the sensitivity of soil water stress signals and PMOC model performance to these common subjective choices (i.e., robustness). Since the goal of this paper is to identify the presence and robustness of soil water stress signals at eddy covariance sites, we will only briefly cover the selected treatments. Section S5 contains further justification on the treatment level selection.

The first four treatments in Table 2 control for different uncertainties in the eddy covariance data on the soil water stress inference.

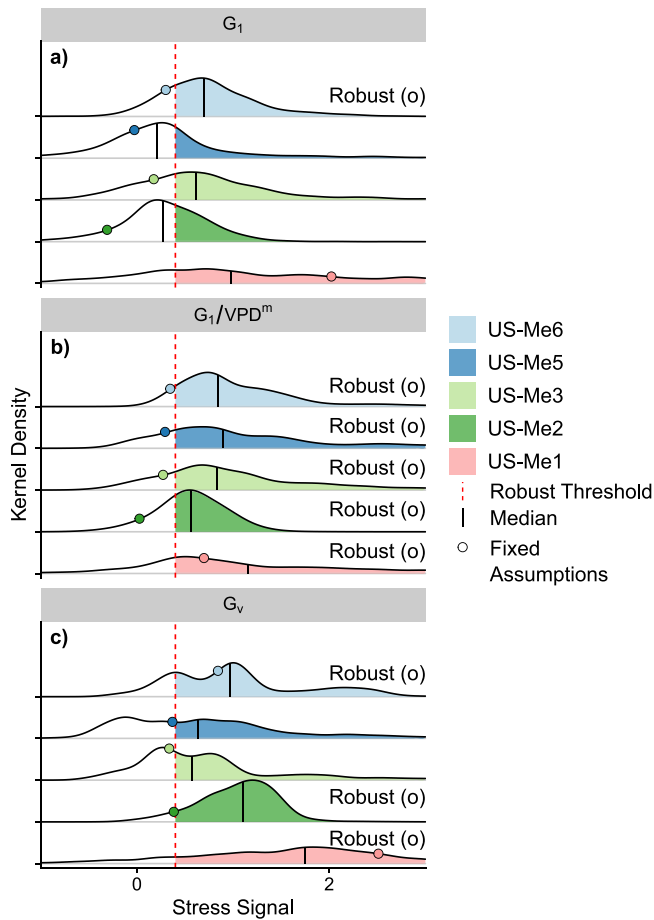


Fig. 3. The kernel density distribution of the (a) G_1 , (b) G_1/VPD^m , and (c) G_v ecosystem soil water stress signals from the 2,304 assumption sets (Table 2) for the five *Metolious ponderosa* pine eddy covariance sites in Oregon. The dashed red line is the critical stress threshold (0.4; see Section 2.3.2), and the filled area under the curve represents the proportion of assumption sets that are deemed to have practically significant soil water stress. The “Robust (O)” label is given to sites that have more than 70% of their assumption sets exceeding the critical stress threshold and is used further in Section 3.3.

Treatment 1 controls for the large, non-normal random uncertainty (i.e., noise illustrated in Fig. 2b) in eddy covariance data (Aubinet et al., 2012) by applying different fit algorithms to the parameter estimation. Treatment 2 controls for the large bias in eddy covariance data (Fig. 2b), attributed to the surface energy budget nonclosure (Mauder et al., 2020), by applying different correction factors. Treatments 3–4 control for the inference uncertainty due to vegetation dynamics (e.g., phenology, plant growth) by either filtering data to a hypothetical “growing season” based on a GPP threshold (black lines in Fig. 2c) or augmenting the data set with MODIS Leaf Area Index (LAI; gray line in Fig. 2c) observations (Myneni et al., 2021). See Section S6 for details on MODIS LAI processing.

The last four treatments in Table 2 control for the inference uncertainties incurred by the PMOC model specification. Treatment 5 (Fig. 2d) controls for uncertainty due to the optimal canopy conductance (i.e., G_c) formulation, while Treatment 6 (Fig. 2e) controls for uncertainties over which PMOC parameters are estimated. Treatment 7 (Fig. 2e) tests the effect of using either G_c derived from observed ET or observed ET itself as the response variable in the parameter estimation. Finally, Treatment 8 (Fig. 2d) tests the effect of derived versus observed VPD inputs into the PMOC model.

We quantify the sensitivity of the inferred soil water stress signals and PMOC model performance to the assumptions in Table 2 using a

digital factorial experiment. The eight treatments containing up to four levels in Table 2 result in 2,304 plausible assumption sets or treatment combinations—thus, soil water stress signals and performance metrics—that a researcher could select in a PMOC analysis. We perform the soil water stress detection discussed in Section 2.1.3 for each unique assumption set (shown in Fig. 2f for two unique assumption sets), which results in 2,304 fits per each θ_p bin (gray points in Fig. 2f–h). We extract the dominant soil water stress signal ($\beta_{1,s,i}$) for each assumption set to create a probability distribution of plausible stress signals for G_1 , G_1/VPD^m , and G_v (Fig. 2i). Likewise, we extract the mean predictive and functional performance metrics corresponding to the soil moisture bins that contained the dominant stress signal ($\beta_{o,p,i}$ and $\beta_{o,f,i}$ in Fig. 2g–h) to create a probability distribution of plausible PMOC model performance (Fig. 2j). We fit the PMOC model in MATLAB using the algorithm specified by Treatment 1 (Table 2) and a 2/3 calibration-validation split. We chose not to quantify sampling variability through methods like bootstrapping (Lin et al., 2018; Bassiouni and Vico, 2021) because we did not want to conflate its variability with the variability due to the assumption choices.

2.3.2. Robust ecosystem soil water stress framework

We propose a robust ecosystem soil water stress framework (or robustness framework) to identify which sites have a detectable soil water stress signal regardless of the selected data and modeling assumptions. Here, we define robustness according to the decision-analysis literature (Mcphail et al., 2018) as the sensitivity of an outcome of interest (i.e., stress signal or performance) to the plausible range of decisions (i.e., data and modeling assumptions). There are numerous metrics to quantify robustness, each with its own trade-offs (Mcphail et al., 2018). Here, we use a satisficing metric known as Starr’s Domain Criteria (Starr, 1963) to quantify robust stress and performance, which calculates the proportion of assumption sets whose stress signals and/or performance metrics exceed a critical threshold.

The critical threshold for the soil water stress signal (i.e., dominant stress signal $\beta_{1,s,1}$ illustrated in Fig. 2f) corresponds to a stress signal large enough to significantly influence ecosystem water and carbon fluxes. Although subjective, we set the threshold to $|\beta_{1,s,i}| \geq 0.4$ (Fig. 2f,i) for the median-normalized soil water stress signal. This threshold corresponds to a 20% drop in the plant parameter (e.g., G_1) over half the soil moisture percentiles (i.e., $0.2/0.5 = 0.4$). Further details on the critical stress threshold are given in Section S3. We will refer to stress signals that exceed this threshold as “practically significant” soil water stress signals throughout this paper. We preliminarily classify the stress at each site using the stress signal pdf median and majority agreement between the ecosystem parameters. We classify a site as “Dry”, “Negligible”, or “Wet” stress if the median lies above, between, or below the positive and negative critical stress thresholds for at least two of the three ecosystem parameters. The few site’s whose ecosystem parameters disagree are labelled “Unsure”.

The critical thresholds for predictive and functional performance metrics correspond to a level of acceptable performance. For predictive performance, we have selected a threshold of $LCE \geq 0.4$ by assuming that r , β_μ , and α in Eq. (3) equal 0.7 (i.e., corresponding to a maximum 30% error), which is reasonable given the large uncertainties in eddy covariance data (see Sect. S7 for more details on derivation of this threshold). For functional performance, we require the $\bar{A}'_{f,T} \geq 0.7$, which represents a maximum 30% deviation between measurements and observations.

These stress signal and performance thresholds are used to calculate the robustness metrics—the proportion of 2,304 assumption sets that exceed the critical thresholds (shown in Fig. 2i and j respectively by the shaded region of the probability density functions (pdfs)). The larger this proportion, the more robust the soil water stress signal and the better the model performance used to infer the signal.

The robustness framework uses these stress and performance robustness metrics to create an overall classification for each eddy covariance

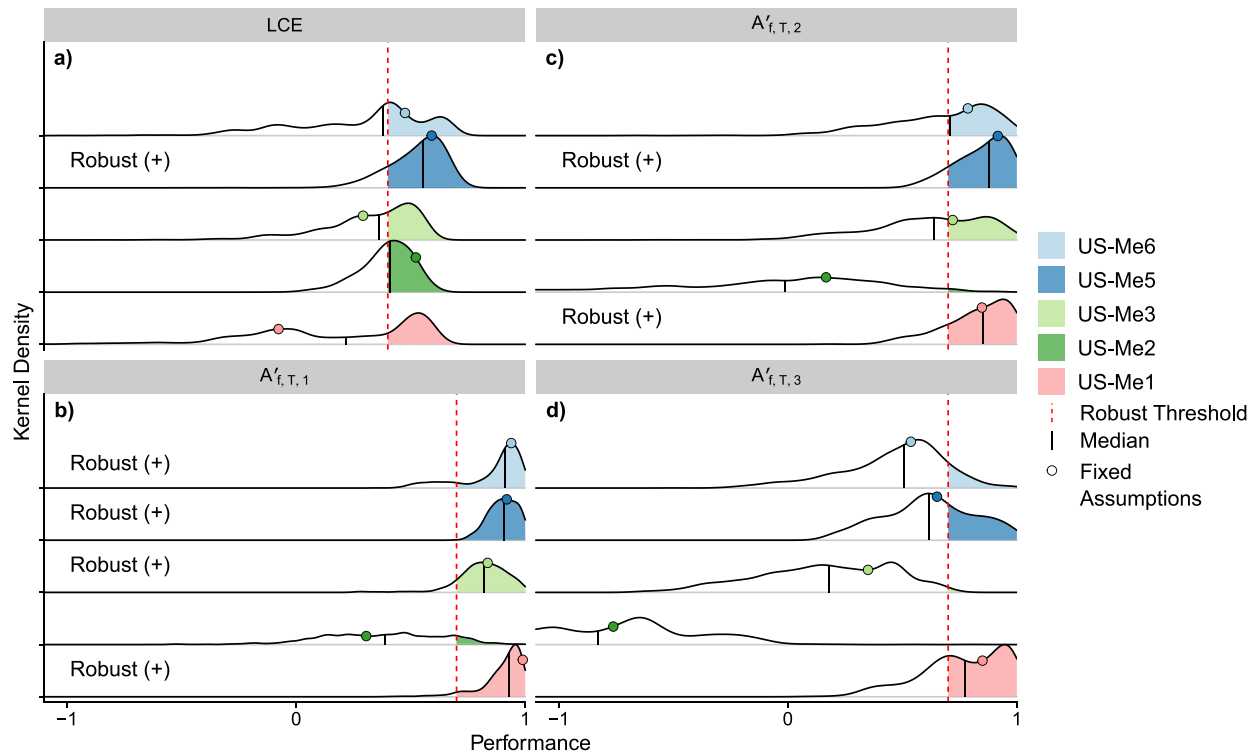


Fig. 4. The kernel density distribution of the PMOC (a) predictive and (b)–(d) functional performance metrics from the 2,304 assumption sets (Table 2) for the five Metolius ponderosa pine eddy covariance sites in Oregon. The dashed red line indicates the critical predictive ($LCE = 0.4$) and functional performance thresholds ($A'_{f,T,j} = 0.7$; see Section 2.3.2), and the filled area under the curve represents the proportion of assumption sets that are deemed to have acceptable performance. The “Robust (+)” label is given to sites that have more than 70% of their assumption sets exceeding the critical performance threshold and is used further in Section 3.3.

site as to whether the site exhibits a robust ecosystem response to soil water stress. Here, we define robustness as 70% of the assumption sets matching the preliminary stress classification (e.g. “Dry”, “Wet”) and/or exceeding the performance thresholds. Note, we are primarily focused on the “Dry” stress classification for this work, but the “Wet” and “Negligible” stress classes do occur. The possible robustness classes for each site are “Stress Only” (○), “Performance Only” (+), or “Stress + Performance” (⊕). The “Stress Only” robustness class indicates that 70% of assumption sets agreed on a “Dry”, “Negligible”, or “Wet” stress classification for at least two of the three ecosystem water use parameters (e.g., a robust “Dry” stress signal shown in Fig. 2i), but the performance metrics failed to meet the robustness criteria (as in Fig. 2j). Therefore, the stress signal may be untrustworthy because the model does not adequately represent the data. Alternately, the “Performance Only” robustness class means that 70% of the assumptions sets had predictive and functional performance metrics jointly exceeding the critical performance thresholds ($LCE \geq 0.4$ and $A'_{f,T} \geq 0.7$, respectively), but the stress signal was too sensitive to the assumptions, which requires closer analysis to estimate the true soil water stress signal. Finally, the “Stress + Performance” robustness class means the stress signals and performance metrics at a site both met the above robustness criteria, and we consider that site to have a robust soil water stress signal. In the example shown in Fig. 2i–j, US-Me2 achieves a robust “Stress Only” class, given its poor model performance. These robustness classes and symbols will be used to rank-order all 151 sites in Section 3.3.

3. Results

3.1. Robust soil water stress inference is a site-specific problem

We first perform a case study on five nearly identical eddy covariance sites in Oregon to illustrate the complexity of robustly inferring

soil water stress signals. These sites are located close to each other (e.g., max distance between sites is 31 km), contain ponderosa pine as the dominant species and share similar climate, weather, and soil textures. More importantly, each site has ample evidence for ecosystem soil water stress given the seasonally dry climate in the eastern Cascade mountains (Irvine et al., 2002; Schwarz et al., 2004; Irvine et al., 2004, 2008; Ruehr et al., 2012). Even so, we can demonstrate that the sensitivity of the soil water stress signals to PMOC assumptions obfuscates stress inference in a site-specific manner.

The diffuse soil water stress signal probability density functions (pdfs) in Fig. 3 signify a lack of robustness. Generally, the median stress signal (black line in Fig. 3) exceeds the critical dry stress threshold for most sites and ecosystem parameters, indicating that the majority of PMOC assumption sets would yield a noticeable dry stress signal (see Section 2.3.2). However, there are still many PMOC assumption sets that yield practically insignificant or even wet soil water stress signals (e.g., unfilled portions left of the red dashed line in Fig. 3). The fact that opposite soil water stress inferences can result from equally valid PMOC assumptions is concerning given that all these Metolius sites have observed periods of dry soil water stress.

Similar to the soil water stress signals, the performance metric pdfs (Fig. 4) show a large spread, implying that many of the assumption sets yield unacceptable performance (unfilled portion of pdf left of the red dashed line in Fig. 4). In particular, the functional performance metrics that involve GPP ($A'_{f,T,2}$ and $A'_{f,T,3}$ in Fig. 4c–d) indicate that the PMOC model struggles to represent the strength of association between GPP and ET (or G_c). As most PMOC studies (Table 1) only select a single assumption set, they do not consider the range of possible performances shown in Fig. 4 nor the implications of these performances on inferring the soil water stress signal. Our results indicate that the inferred soil water stress signals should be qualified by their corresponding performance metrics.

The differing pdf shapes across sites in Figs. 3–4 indicate that the robustness of soil water stress signals and model performance are site-specific. To further illustrate this point, we have plotted the inferred stress signals and performance metrics for an arbitrarily chosen set of fixed PMOC assumptions as filled circles in Figs. 3–4. At each site, the circle lies at a different percentile of the pdf, indicating that the fixed assumption set uniquely affects inference at each site. For example, the fixed assumption set for US-Me1 in Fig. 3 causes a higher percentile stress signal than at the other four sites. Our preliminary analysis indicates that the site-specific robustness results are a consequence of unique site characteristics, such as stand age, disturbance history, and observation period (see Sect. S8). Regardless, the site-specific robustness results indicate that most PMOC studies—that use a single assumption set for consistency across sites—may actually have site-specific biases in their inferred stress signals that complicate comparisons across sites and studies.

3.2. Lack of robustness confounds generalizing ecosystem soil water stress signals

The case study results (Section 3.1) raise concerns over the common practice in PMOC studies (Table 1) of aggregating inferred soil water stress signals into broad ecosystem categories to draw general conclusions about ecosystem water use strategies. We hypothesize that site-specific, non-robust soil water stress signals and model performance are prevalent at sites outside of our case study, which will compromise the ability to detect patterns between broad ecosystem classes. To test this hypothesis, we perform a typical multi-site analysis for 151 eddy covariance sites that groups inferred soil water stress signals for a single assumption set (i.e., the median of each site's pdf) into broad ecosystem categories defined by the IGBP plant functional types (PFTs) and the annual Dryness Index (DI; see Section 2.1.4 for details). We test if the lack of robustness in soil water stress signals and model performance obfuscate expected trends across ecosystem categories.

Fig. 5 shows that the intra-ecosystem variability in the median stress signals is far greater than the inter-ecosystem variability with respect to DI or PFTs. Previous studies indicate that sites with higher DI correspond to larger dry soil water stress signals given the presence of seasonal water stress (Novick et al., 2016; Fu et al., 2022a) or that grass/annual sites (GRA or CRO) would have more aggressive water use strategies than forested sites (DBF or ENF) given their shorter life-span (Lin et al., 2015). However, the intra-ecosystem variability in Fig. 5 overwhelms any statistically significant differences between ecosystem categories (see Sect. S9 for details).

These inconclusive inter-ecosystem patterns are further reinforced by the lack of robustness in the soil water stress signals to the PMOC assumptions at most sites (Fig. 5b). The low signal-to-noise ratio (SNR; Fig. 5b; analogous to high spread in Fig. 3) indicates that a researcher could infer a different stress signal (e.g. wet stress versus dry stress) based on subjective assumptions. Furthermore, as in the case study, the robustness of the soil water stress signal is site-specific (Fig. S7), meaning that using a fixed assumption set may result in a site-specific bias from the true ecosystem soil water stress signal, which could contribute to the large intra-ecosystem variability shown in Fig. 5a.

Limitations of the PMOC inference approach also contribute to the inability to generalize ecosystem water use strategies. The commonly-used ecosystem water use parameter G_1 struggles to detect soil water stress as indicated by fewer practically significant stress signals (number of points exceeding red dashed line in Fig. 5a tallied in Fig. S5b) and lower SNR (Fig. 5b) across ecosystem categories compared to G_1/VPD^m and G_v . We discuss this point further in Section 4.1. The poorer predictive (LCE) and functional performance ($\overline{A'}_{f,T}$) of the PMOC model (Fig. 6a) at forested sites (DBF and ENF) compared to annual/grass sites (CRO and GRA) may also increase the level of intra-ecosystem variability in Fig. 6a through inaccurate parameter

estimation. These PMOC performance differences between PFTs appear robust (high SNR in Fig. 6b) and are discussed in Section 4.2.

Finally, a practical explanation for the failure of the ecosystem stress generalization is the inclusion of eddy covariance sites that have not observed periods of soil water stress—or, at least, not signals detectable by the PMOC analysis. Unlike the Metolius case study in Section 3.1, we do not know how many of the 151 eddy covariance sites actually observed soil water stress. Only 35%–57% (5%–12%) of the 151 eddy covariance sites exhibited practically significant dry (wet) soil water stress based on their median soil water stress signal (Fig. S5b), depending on the plant parameter (Fig. S5a). Many eddy covariance sites are hydric or mesic (Fig. 1) and have less than 10 years of observations (Fig. S6), meaning that there is a good chance that the observation period did not experience an intermittent drought.

3.3. Ranking ecosystem soil water stress signals by robustness

The soil water stress signals inferred from numerous eddy covariance sites are not robust to the subjective PMOC assumptions, undermining meaningful conclusions about ecosystem responses to soil water stress. In this section, we apply the robust ecosystem soil water stress framework (Section 2.3.2) to identify which sites have a robust soil water stress signal as well as general deficiencies in the PMOC inference approach. The robustness framework classifies and ranks each site by its ability to yield consistent soil water stress signals and acceptable model performance with respect to the numerous PMOC assumption sets (Fig. 2i–j). We visualize the robustness framework results with heat maps for all 151 eddy covariance sites (Figs. 7–8). We will first examine these robustness results for the Metolius case study sites (Section 3.1) to orient the reader to the heat maps, followed by an analysis of all 151 eddy covariance sites.

The robustness framework heat maps (Figs. 7–8) encapsulate the key information from the stress signal and performance metric pdfs (e.g., Figs. 3–4) into three columns: overall robust stress classification (column 1), stress signal robustness (column 2), and performance metric robustness (column 3). All five Metolius case study sites (rows with black arrows in Fig. 7) indicate a dry soil water stress signal (red in column 1) given that their median stress signals exceed the critical stress threshold of 0.4 for two out of three plant parameters (Fig. 3a–c). However, the robustness and, thus, ranking of each site's stress signal varies as shown by the differing symbols in column 1 of Fig. 7.

US-Me5 ranks the highest with robust performance only (“+” in column 1), as more than 70% of the runs jointly exceed the predictive ($LCE \geq 0.4$) and functional performance threshold ($\overline{A'}_{f,T} \geq 0.7$). The robust performance classification is based on the dark green and “+” symbols in column 3 that summarize the median and robustness metrics, respectively, of the individual performance metric pdfs in Fig. 4. Note that $\overline{A'}_{f,T,3}$ does not individually meet the robustness threshold (no “+” in column 3 of Fig. 7, and no “Robust (+)” label in Fig. 4d), but $\overline{A'}_{f,T}$ does (see Section 2.2). Unfortunately, US-Me5 does not have a robust stress signal as only G_1/VPD^m indicates robustness (“○” in column 2, and “Robust (○)” label Fig. 3b) while G_1 and G_v have too many assumption sets yielding practically insignificant soil water stress (Fig. 3a,c). Alternately, US-Me1, US-Me6, and US-Me2 are classified as robust stress only (“○” in column 1), indicating that two of the three ecosystem parameters have at least 70% of assumption sets indicating dry soil water stress (“○” in column 2). These three sites are ranked below US-Me5 because they have very poor predictive performance (no “+” in LCE sub-column of Fig. 7 and no “Robust (+)” label in Fig. 4a), indicating that the robust dry soil water stress signal may be dubious due to unacceptable PMOC model performance. Lastly, US-Me3 receives no robustness class (no symbol in column 1), which means the soil water stress signal and performance are too sensitive to the PMOC assumptions.

Although all five Metolius sites have experienced soil water stress (see Section 3.1), we could not classify any of them as having robust

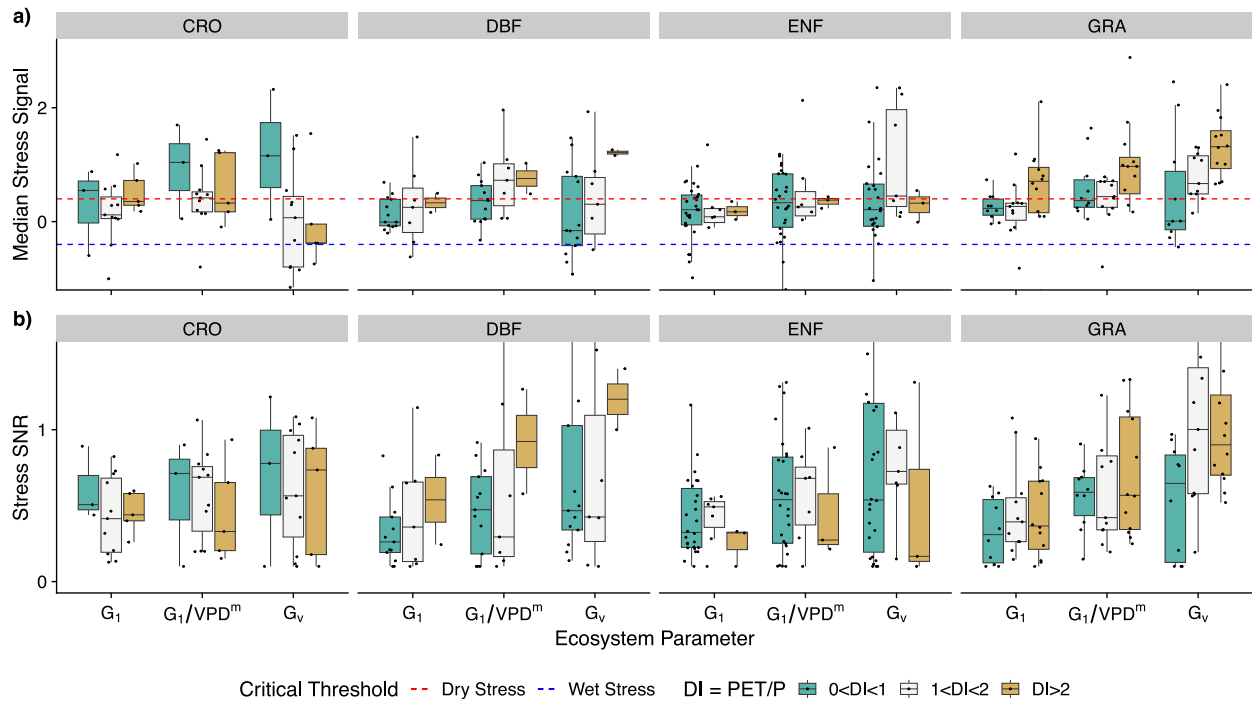


Fig. 5. The soil water stress signal pdf (a) median and (b) signal-to-noise ratios (SNR = median/interquartile range) for eddy covariance sites grouped by plant functional type (PFT) and annual Dryness Index (see Section 2.1.4). Here, we show only the most prevalent PFTs (110 of 151 sites; see Figure S4 for all PFTs): cropland (CRO), deciduous broadleaf forest (DBF), evergreen needleleaf forest (ENF), and grassland (GRA). The black markers are the site-specific values jittered horizontally for visibility. In (a), the points above (below) the red (blue) dashed line have a dry (wet) soil water stress signal as the median of their stress signal pdf (e.g., black line in Fig. 3) exceeds the critical stress threshold. In (b), a larger value indicates less sensitivity to the PMOC model assumptions.

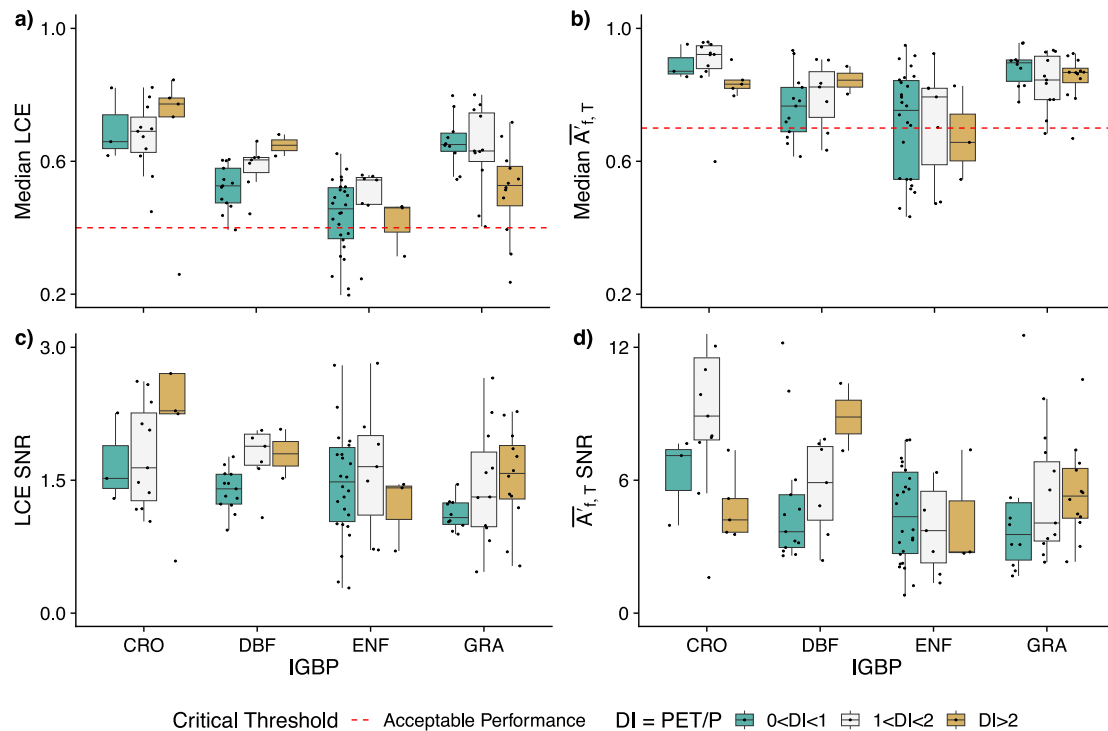


Fig. 6. The (a) predictive and (b) functional performance metric pdf medians and their corresponding signal-to-noise ratios (SNR) in (c)–(d) for the same ecosystem categories in Fig. 5. The black markers are the site-specific median performance metrics jittered horizontally for visibility. Values in (a)–(b) above the red dashed line have acceptable performance according to the median of their performance pdf (e.g., black line in Fig. 4), while higher values in (c)–(d) indicate less sensitivity to the PMOC model assumptions. See caption of Fig. 5 for abbreviation descriptions. See Fig. S8 in the supplement for all PFTs.

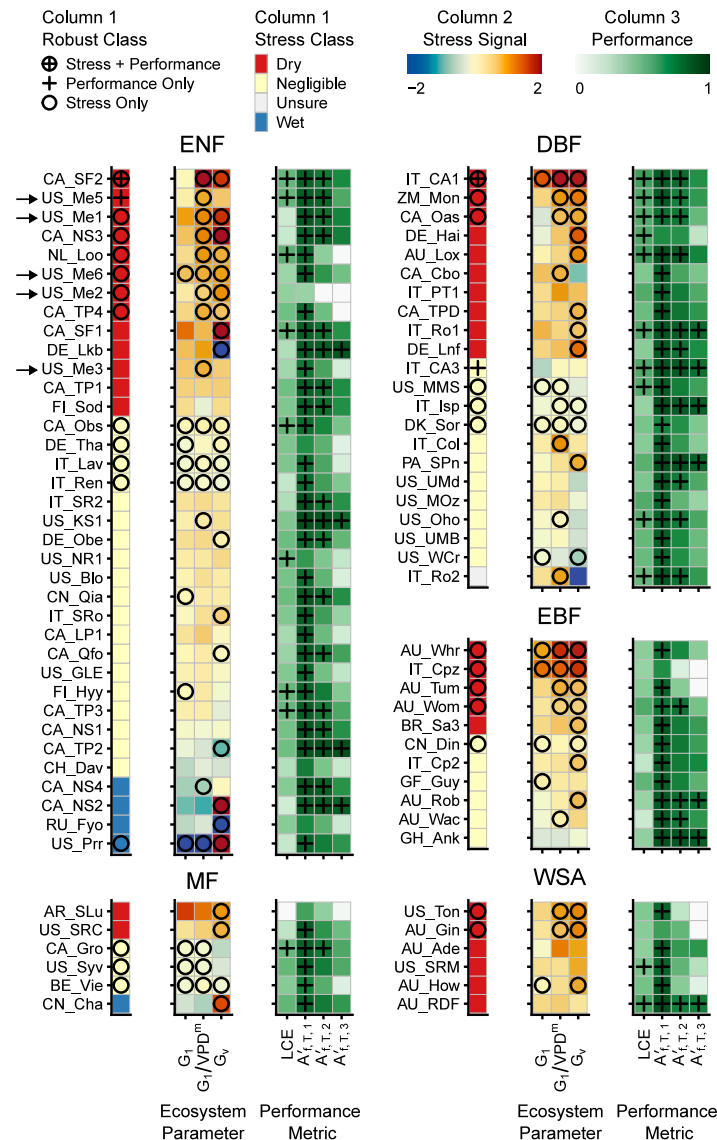


Fig. 7. The robust ecosystem soil water stress framework (see Section 2.3.2) heat maps for all forested eddy covariance sites, split up by PFTs: Evergreen Needleleaf Forest (ENF), Mixed Forest (MF), Deciduous Broadleaf Forest (DBF), Evergreen Broadleaf Forest (EBF), and Woody Savanna (WSA). Column 1 classifies the soil water stress class (color) and corresponding robustness class (symbol) discussed in Section 2.3.2 and illustrated in Fig. 2i–j. Column 2 shows the median stress signal and the corresponding stress robustness class (Fig. 2i). A black “○” (“+”) indicates that the 70% stress (performance) robustness criteria was met for an ecosystem parameter (performance metric). A “⊕” in the Column 1 indicates the site has fulfilled both the robust stress and performance criteria, and has a robust soil water stress signal. Note, sites can have robust negligible soil water stress, e.g., CA_Obs.

stress and performance (“⊕” in column 1 of Fig. 7). Unfortunately, this lack of robustness in both stress and performance appears pervasive across all 151 eddy covariance sites as only 7 sites have robust stress and performance (“⊕” in column 1 of Figs. 7–8), while 7 sites have robust performance only and 48 sites have robust stress only (32 of these are dry soil water stress). This indicates that between 5% (“⊕” in column 1) and 36% (“+” or “○” in column 1) of the 151 sites have a robust soil water stress signal, depending on whether we consider model performance. Furthermore, the varying magnitudes of stress signals and performance (colors indicating median value in columns 2–3) as well as their robustness classes (symbols in column 1) are poorly described by PFT groups in Figs. 7–8, reinforcing the site-specific nature of robust soil water stress inference discussed in Sections 3.1–3.2.

The ecosystem water use parameter G_1 , primarily used to assess soil water stress in PMOC studies (Table 1), is rarely robust (few “○” in G_1 sub-column of Figs. 7–8). In these cases, the parameters G_1/VPD^m and G_v must both exceed our robust stress threshold to earn

a robust classification (e.g., US-Me1, US-Me2, CA-SF2 in Figs. 7, and 3a,c). This two-out-of-three rule (see Section 2.3.2) allows more flexible detection of soil water stress at the risk of misattributing seasonality in environmental forcings as soil water stress (see Section 4.1 for discussion). Without G_1/VPD^m and G_v , we would only have 4 sites (IT-CA1/DBF, US-Arb/GRA, DE-Seh/CRO, and AU-ASM/SAV) out of 151 classified as robust soil water stress and performance.

The PMOC model performance appears to be the primary limitation to robustly identifying soil water stress given there are only 14 sites with robust performance (“⊕” or “+” in column 1) compared to 55 sites with robust stress (“⊕” or “○” in column 1). More specifically, the predictive performance (LCE) is the most limiting, followed by the two functional performance metrics that include GPP ($A'_{f,T,2}$ and $A'_{f,T,3}$). Generally speaking, the grass/annual vegetation sites (Fig. 8) have higher predictive and functional performance (darker colors in column 3 of Figs. 8 over 7), indicating that the PMOC model may more adequately describe these systems. However, the LCE is still

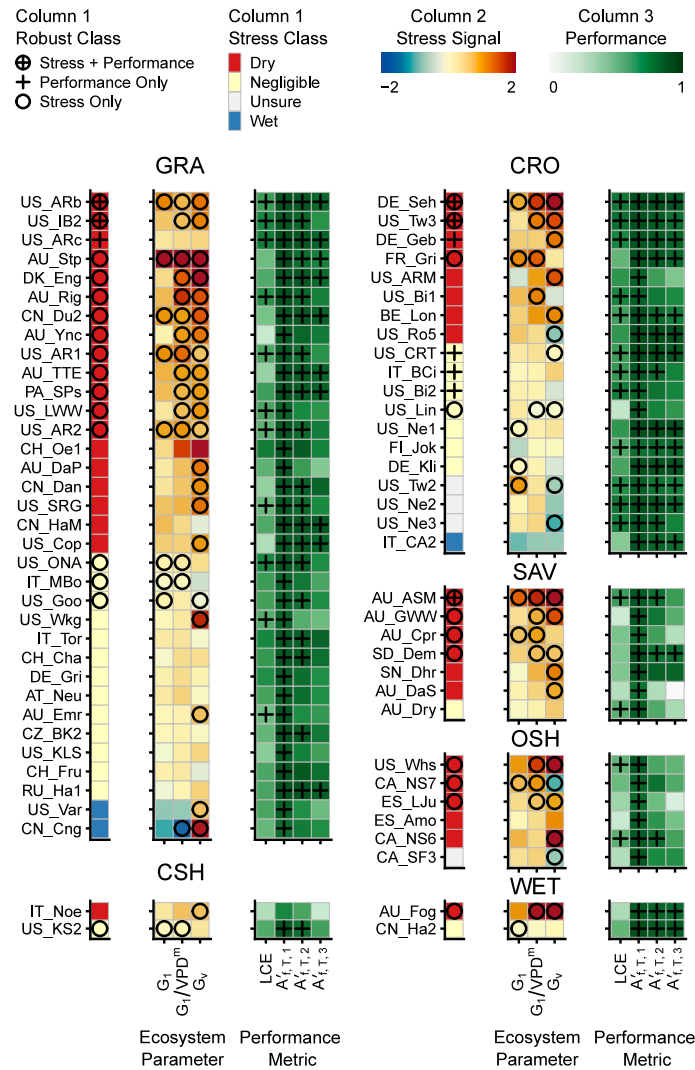


Fig. 8. Same as Fig. 7, except for all grass/annual ecosystem PFTs: Grassland (GRA), Closed Shrubland (CSH), Cropland (CRO), Savanna (SAV), Open Shrubland (OSH), and Wetland (WET). Note, sites can have robust negligible soil water stress, e.g., US_ONA.

not robust in many cases (i.e., limited “+” in LCE sub-column of column 3), indicating large predictive performance sensitivity to PMOC assumptions and/or high levels of noise in the eddy covariance data (see Section 4.1). For functional performance, the prevalence of poor $A'_{f,T,2}$ and/or $A'_{f,T,3}$ values (last two sub-columns in column 3) may indicate large sensitivity to the growing season or LAI assumptions (Treatments 3–4 in Table 2) or the inability of the PMOC model to represent the influence of GPP on ET (e.g., non-stomatal limitations Zhou et al., 2013). Regardless, the inadequate performance of the PMOC model must be resolved to understand which eddy covariance sites have robust soil water stress signals.

4. Discussion

Our proposed robustness framework quantifies the uncertainties associated with inferring ecosystem soil water stress signals from eddy covariance data, stemming from the numerous data uncertainty and model specification assumptions. We find that these stress signals are often not robust, and caution against using a single assumption set for PMOC and similar inference approaches. The lack of robustness is likely due to both practical identifiability of the ecosystem water use strategy parameter G_1 and poor PMOC model performance. The

non-robust stress signals also hinder generalization of ecosystem water use strategies using common groupings of PFTs and climatic indices. Nevertheless, we find that grass/annual vegetation sites tend to have more robust performance compared to forested sites. We will now discuss the implications of these findings for inferring ecosystem soil water stress from eddy covariance data.

4.1. Lack of robustness in the soil water stress signals

The small number of sites with both robust soil water stress signals and model performance metrics (only 7 sites total with “⊕” in column 1 of Figs. 7–8) with respect to the plausible PMOC assumptions calls for greater care when inferring ecosystem water use strategies from eddy covariance data with PMOC or similar inference approaches (e.g., statistical models). Currently, many studies (Table 1) select a single set of PMOC assumptions, or, in the best case, explicitly test the sensitivity of several assumptions in a one-at-a-time fashion (Knauer et al., 2018; Lin et al., 2018). Unlike the factorial approach used in this study, the one-at-a-time sensitivity approach likely underestimates the assumption uncertainty (i.e., overestimates robustness) by ignoring the interactions between assumptions (Table 2). Our results reveal that

different assumption sets can yield diametrically opposed stress conclusions (e.g., wet and dry stress signals possible for all US-Me sites in Fig. 3) and/or differing model performances (e.g., Fig. 4). Furthermore, we illustrate that a consistent set of PMOC assumptions will not provide a consistently biased estimate of the true underlying soil water stress signal across sites (e.g., the filled circles in Figs. 3–4). In other words, the sensitivity of the inferred stress signal and model performance to a set of PMOC assumptions appears site-specific. Therefore, comparing results across studies that use different PMOC assumptions should be avoided. Future work will focus on understanding which assumptions maximize model performance and stress signal robustness for each site, and determine if these best assumption sets can be generalized across sites (Sloan, 2023).

Our results highlight that G_1 is often not robust to PMOC assumptions (few “O” in G_1 sub-column of Figs. 7–8), despite its widespread use (Table 1) and association with ecosystem water use strategies (see Section 2.1.3). The decision of which PMOC parameters to fit (Treatment 6 in Tables 1–2) may be the most influential for determining G_1 robustness. Given the high level of noise in eddy covariance data (e.g., Fig. 2b), the parameters G_1 , G_o , m may not be practically identifiable (Guillaume et al., 2019) and correlations between parameters during the fitting process can yield unrealistic G_1 values. Previous PMOC studies have avoided the identifiability issue by simply fitting G_1 (Medlyn et al., 2017; Knauer et al., 2018), while some studies argue that fitting the other parameters quantifies important physical processes, such as soil evaporation (G_o) and sub-optimal plant sensitivity to VPD (m) (Novick et al., 2016; Lin et al., 2018; Li et al., 2019). In this study, we supplemented the poorly-constrained G_1 estimates by examining two additional ecosystem parameters, G_1/VPD^m and G_v . Nevertheless, using these parameters runs the risk of masquerading environmental seasonality (e.g., in net radiation, temperature, VPD, and leaf area) as ecosystem response to soil water stress. Disentangling the multi-scale coupling of these environmental forcings on ecosystem soil water stress response is an area of active research (Novick et al., 2019; Feldman et al., 2021; Fu et al., 2022a), requiring data-driven and mechanistic approaches. More detailed parameter identifiability (Guillaume et al., 2019) and targeted data analysis (Fu et al., 2022b) are required to ensure that inferred G_1 values represent the ecosystem’s water use strategy and are not artifacts of correlations with other parameters or environmental variables.

Poor model performance is the primary constraint to robust ecosystem soil water stress inference (only 14 of 151 sites have robust performance in Figs. 7–8). The predictive performance (LCE) appears the most limiting (fewer “+” in the LCE sub-column of Figs. 7–8), which could stem from limitations in the eddy covariance data quality. The inherent noise and bias in eddy covariance data (Fig. 2b and Sect. S5) places an upper bound on achievable model performance, which we have estimated for each site (see Sect. S6). The performance upper bounds appear site-specific with many sites near or below the acceptable predictive performance threshold (Fig. S9). The site-specific performance upper bound may also explain part of the site-specific robustness of ecosystem soil water stress signals, as sites with lower upper bounds (i.e., more noise) would likely not meet the acceptable performance threshold, and may also have less robust stress signals due to greater G_1 identifiability issues. These data quality issues may also be exacerbated by the extensive filtering required by the PMOC analysis (Fig. 2a), which, on average, removes over 90% of the available data (Fig. S10 in supplement), leaving less information to inform the PMOC model parameters. Overall, the performance deficiencies motivate future work (1) identifying which PMOC assumptions adversely affect performance (Sloan, 2023), and (2) testing more complex interpretable model formulations that can leverage more of the eddy covariance data.

4.2. Soil water stress cannot be generalized across ecosystem categories

Our results challenge the validity of grouping site-level soil water stress signals inferred from eddy covariance data by broad ecosystem categories to make general conclusions about ecosystem water use strategies under soil water stress. The ecosystem categories do not explain variations in the median stress signal behavior (Fig. 5a and Sect. S9), which is further confounded by the lack of robustness to PMOC assumptions (low SNR in Fig. 5b). The Metolious case study (Section 3.1) and robustness framework heat maps (Figs. 7–8) emphasize the site-specific nature of ecosystem soil water stress inference. We hypothesize that the ecosystem generalization of soil water stress fails due to both the complexity in ecosystem water use strategies and practical data limitations.

Plant water use strategies emerge from the interplay of plant hydraulic and functional traits with their environment over time. These plant traits exhibit wide inter- and intra-species variability (Anderegg, 2015) and plasticity to environmental forcings (e.g., increased root growth after a drought Rowland et al., 2023), hindering generalization by broad categories (Matheny et al., 2017; Kannenberg et al., 2021). The difficulty in generalizing water use strategies compounds when scaling from plant to ecosystem, as multiple species may have differing water use strategies that are not easily captured by the broad categories used here (PFT and DI) and in other PMOC studies (Table 1). The Metolious case study (Section 3.1) provides evidence for different ecosystem water use strategies under the same ecosystem classification (all sites are hydric ENF) driven by site-specific differences in stand age, disturbance history and observation period (see Sect. S8 for further discussion).

The eddy covariance observation periods themselves also contribute to the failure of ecosystem soil water stress generalization, as most sites have less than 10 years of observations (c. 75% in Fig. S6). Our results estimate that 5%–21% (ignoring 16 sites with negligible robust stress only in Figs. 7–8) of the sites experienced dry soil water stress based on the robustness heat maps (Figs. 7–8). Therefore, many sites either do not observe soil water stress during the short observation period or the PMOC model cannot detect the signal due to the lack of robustness. The strongest dry soil water stress signals are found at xeric sites (Figs. 5a and 8)—primarily ecosystems with shorter vegetation (e.g., GRA, CRO, SAV)—as these sites have soil water stress seasonally, making detection more likely. Unfortunately, the stress signals from the larger number of mesic and hydric sites (Fig. 1) are more ambiguous, being more indicative of whether a site observed a period of soil water stress rather than the true response of that ecosystem to soil water stress. The poor performance of ecosystem categories may also be explained by the fact that the annual DI poorly represents dryness over short observation periods or in seasonal environments (e.g., seasonal droughts in Mediterranean climates; Feng et al. (2019)). However, we did test other ecosystem categories using a tower-measured DI and a seasonality index, but the intra-ecosystem variability still vastly outweighed inter-ecosystem variability (Fig. S11–S12). Thus, we would not expect any similarly broad ecosystem categories to detect patterns in stress signals from hydric or mesic sites. We recommend more thoroughly characterizing soil water stress signals at the site-level before attempting to look for patterns between numerous sites.

4.3. PMOC models perform better at grass/annual ecosystems

Our results do indicate better PMOC model performance for grasses/annual (e.g., crops, grasslands) sites compared to those with taller trees. The median functional and predictive performance metrics were higher for GRA and CRO ecosystems compared to ENF and DBF (Fig. 6a–b), while robustness heat maps show a greater number of sites with robust performance (more “+” markers in Fig. 8 compared to Fig. 7). The superior performance likely reflects the reduced complexity of grass/annual vegetation stress responses compared to those of

forested ecosystems. We will now discuss several factors that favor performance improvements in grass/annual versus forested vegetation.

The PMOC approach here primarily uses shallow soil moisture sensors, which are likely more representative of root zone moisture dynamics in grass/annual vegetation compared to deeper-rooted trees. For instance, at US-Me2, nearly half the root water uptake occurs below 80 cm (Irvine et al., 2004). Other studies have attempted to counter this by using deeper moisture sensors where available (Novick et al., 2016; Bassiouni and Vico, 2021); however, many sites do not have soil moisture profiles with a long observation period (US-Me2 is an exception), and deeper soil moisture observations may not represent the surface soil evaporation component of the PMOC model (i.e., fitting G_o).

The poorer PMOC model performance for forested compared to grass/annual vegetation sites may also be due to the omission of plant hydraulic transport. Sloan et al. (2021) found that the coupling of soil and atmospheric water stress through a Plant Hydraulics Model (PHM) was critical for carbon, water, and energy flux predictions at US-Me2, which has been supported by similar studies mechanistically modeling a range of eddy covariance sites (Kennedy et al., 2019; Eller et al., 2020; Sabot et al., 2020). The inability of the PMOC model to realistically respond to VPD and soil moisture could explain both the poor LCE and $A'_{f,T,2} - A'_{f,T,3}$ values at forested sites (column 3 in Fig. 7). Allowing variation in G_1 and/or m with respect to VPD (in addition to soil moisture) may better represent plant hydraulic behavior and improve the predictive and functional performance at forested sites.

Finally, the PMOC performance discrepancy may be due to the differences in long-term dynamics between annual and perennial vegetation. For grass/annual sites, the vegetation resets every year, leading to similar ecosystem water use strategies (especially in cropped systems) as long as there are no major disturbances or changes in plant species. At perennial forested sites, the exhaustion of carbon pools and changes in allocation between roots, stems and leaves due to drought will result in plasticity of underlying plant hydraulic traits controlling the ecosystem water use strategy over years (Anderegg et al., 2015; Rowland et al., 2023). For instance, the data from US-Me2 covers nearly 20 years, and it is well known that this site responds non-linearly to single versus multi-year droughts (Thomas et al., 2009). These dynamics will be missed by fitting time-constant parameters as we have in this study. Knauer et al. (2018) found inter-annual variability of G_1 derived from six forested eddy covariance sites under well-watered conditions, and we suggest future work perform a similar analysis at more sites under soil water stress.

4.4. Limitations

A primary limitation of this study is the subjective nature of the robustness criteria (Section 2.3.2). We used a 30% threshold for error for performance ($A'_{f,T} \geq 0.7$ and $LCE \geq 0.4$) and the robust satisficing criteria (>70% of assumption runs must meet slope and performance criteria). For error, 30% seems like a practical upper bound, as higher error values may be undesirable. Thus, our estimates represent an upper bound on how many sites have robust performance, as a lower acceptable error would result in fewer sites able to meet the criteria. We adjusted the robust satisficing criteria and the critical stress threshold and found that higher (lower) values reduced (increased) the number sites with robust classifications, but the robustness rank-order of sites does not change drastically (Figure S13). Additionally, the robustness metric itself creates some non-intuitive results as the performance metrics appear more robust in Fig. 6c–d (i.e., higher SNR), yet only 14 of 151 sites have robust performance (“⊕” or “+” in column 1 of Figs. 7–8). Although, the performance metrics are less sensitive to assumptions than the stress signals, their values fail to meet the critical performance thresholds for many assumption sets. Further refinements to the robust ecosystem stress—such as using variance-based robustness

metrics (Mcphail et al., 2018)—could help ensure that our robustness framework is robust to our assumptions.

The stress signal extraction (Section 2.1.3) required assumptions of (segmented) linearity, which may provide inaccurate stress signals for non-linear stress responses. Furthermore, binning by soil moisture percentiles rather than parametrizing the soil moisture stress response on G_o , G_1 , and m can mask non-linearity in the stress response and reduces the number of data points for the linear regression (i.e., max of 10 points illustrated in Fig. 2f). However, binning by soil moisture is a common approach in many PMOC studies (Table 1), and allowed testing different functional forms for the stress response. Regardless, this analysis could be repeated without binning and specifying a piecewise linear or other function multiplied by G_1 in the PMOC model. Finally, the choice to use soil moisture percentiles (θ_p , Novick et al., 2016; Lin et al., 2018; Li et al., 2019) to represent soil moisture rather than the observed volumetric water content (θ , Bassiouni et al. (2018) and Fu et al. (2022a)) or estimating soil water potential with a pedotransfer function (Bassiouni and Vico, 2021) could affect the results. We selected θ_p because it (1) standardizes results between sites with differing θ ranges, (2) maps closer to soil water potential than θ (Novick et al., 2016), (3) avoids large uncertainties in pedotransfer functions (Novick et al., 2022), and (4) contextualizes soil moisture in terms of frequency of occurrence at a site. However, future refinements to the robustness framework could test a change in soil moisture variable. Lastly, we were not able to test all pertinent assumptions for the PMOC and similar modeling approaches (Table 1), as the number of simulations grow exponentially with additional treatment levels. Similar to other PMOC studies, we assumed baseline data filtering assumptions (see Section 2.1.1), and used the recommended atmospheric conductance (G_a) parametrization from Knauer et al. (2018) (see Sect. S2). We also were not able to test the effects of deeper soil moisture sensors as many sites do not have profile measurements with adequate coverage. Finally, we restricted the analysis to PMOC models over simpler data-driven models (e.g., Koster et al. (2009)) because the PMOC models tend to be more interpretable with respect to stomatal closure in response to soil water stress. The PMOC model explicitly attempts to control for other environmental variables that influence stomatal closure and are correlated with soil moisture (e.g., R_n , VPD , GPP , T_a) when inferring the soil water stress signal. Simpler data-driven approaches are useful for separating regimes of energy-limited versus water-limited ET (Koster et al., 2009), but eventually have to average over longer timescales (e.g., daily, seasonally) (Koster et al., 2009; Fu et al., 2022a) to control for other drivers of stomatal closure. However, future refinements of the robustness framework could assess these simpler, data-driven models.

5. Conclusions

Our results are a first attempt to comprehensively identify which eddy covariance sites have observed soil water stress and assess the current ability to infer these soil water stress signals. We find that most ecosystem soil water stress signals inferred from eddy covariance data are not robust with respect to the numerous data and modeling assumptions required by current inference approaches. Furthermore, the robustness of the soil water stress signals and model performance are site-specific, which undermines current practices of using a single assumption set and generalizing site-level results by broad ecosystem categories. Our proposed robust ecosystem soil water stress framework provides a road map for quantifying and ranking the site-specific robustness of stress signals inferred from eddy covariance data that can be extended to additional sites and updated data sets. The robustness framework revealed that the deficient interpretable model performance and poorly constrained ecosystem parameters hinder robust soil water stress inference. These results will guide future research in identifying the assumption sets that maximize model performance and minimize uncertainty in the inferred soil water stress signal. Only once soil water stress can be robustly inferred from eddy covariance data will we be able to confidently extract meaningful conclusions about ecosystem responses to soil water stress around the globe.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The relevant codes for this research are freely available at <https://github.com/sloan091/afm-robust-stress-id>. The data used in this research are freely available from the FLUXNET2015 (<https://fluxnet.org/data/fluxnet2015-dataset/>), AmeriFlux (<https://ameriflux.lbl.gov/data/download-data/>), and MODIS (https://doi.org/10.5067/MODIS/MC_D15A2H.061) databases.

Acknowledgments

B.P.S and X. F. acknowledge the support from the National Science Foundation CAREER award DEB-2045610. B.P.S would also like to acknowledge support from Oak Ridge National Laboratory. B.P.S and X.F. are also grateful to the two anonymous reviewers for their helpful suggestions. This work used global eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization were carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center and the OzFlux, ChinaFlux, and AsiaFlux offices. Funding for the AmeriFlux data portal was provided by the U.S. Department of Energy Office of Science.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.agrformet.2023.109744>.

References

- Anderegg, W.R.L., 2015. Minireview Spatial and temporal variation in plant hydraulic traits and their relevance for climate change impacts on vegetation. *New Phytol.* 205, 1008–1014. <http://dx.doi.org/10.1111/nph.12907>, URL: www.newphytologist.com.
- Anderegg, W.R., Schwalm, C., Biondi, F., Camarero, J.J., Koch, G., Litvak, M., Ogle, K., Shaw, J.D., Shevliakova, E., Williams, A.P., Wolf, A., Ziaco, E., Pacala, S., 2015. Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science* 349 (6247), 528–532. <http://dx.doi.org/10.1126/science.aab1833>, URL: <https://www.science.org/doi/10.1126/science.aab1833>.
- Aubinet, M., Vesala, T., Papale, D., 2012. Eddy Covariance: A Practical Guide to Measurement and Data Analysis. Springer Netherlands, URL: <https://doi.org/10.1007/978-94-007-2351-1>.
- Bassiouni, M., Higgins, C.W., Still, C.J., Good, S.P., 2018. Probabilistic inference of ecohydrological parameters using observations from point to satellite scales. *Hydrol. Earth Syst. Sci.* 22 (6), 3229–3243. <http://dx.doi.org/10.5194/hess-22-3229-2018>, Publisher: Copernicus GmbH.
- Bassiouni, M., Vico, G., 2021. Parsimony vs predictive and functional performance of three stomatal optimization principles in a big-leaf framework. *New Phytol.* 231 (2), 586–600. <http://dx.doi.org/10.1111/NPH.17392>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1111/nph.17392>. Publisher: John Wiley & Sons, Ltd.
- Boese, S., Jung, M., Carvalhais, N., Teuling, A.J., Reichstein, M., 2019. Carbon-water flux coupling under progressive drought. *Biogeosciences* 16 (13), 2557–2572. <http://dx.doi.org/10.5194/bg-16-2557-2019>, Publisher: Copernicus GmbH.
- Cowan, I.R., Farquhar, G.D., 1977. Stomatal function in relation to leaf metabolism and environment: Stomatal function in the regulation of gas exchange. In: *Integration in Activity of the Higher Plant*. pp. 471–505.
- Dewar, R., Mauranen, A., Mäkelä, A., Hölttä, T., Medlyn, B., Vesala, T., 2018. New insights into the covariation of stomatal, mesophyll and hydraulic conductances from optimization models incorporating nonstomatal limitations to photosynthesis. *New Phytol.* 217 (2), 571–585. <http://dx.doi.org/10.1111/nph.14848>, URL: www.newphytologist.com.

- Drake, J.E., Power, S.A., Duursma, R.A., Medlyn, B.E., Aspinwall, M.J., Choat, B., Creek, D., Eamus, D., Maier, C., Pfautsch, S., Smith, R.A., Tjoelker, M.G., Tissue, D.T., 2017. Stomatal and non-stomatal limitations of photosynthesis for four tree species under drought: A comparison of model formulations. *Agricult. Forest Meteorol.* 247, 454–466. <http://dx.doi.org/10.1016/j.agrformet.2017.08.026>, Publisher: Elsevier.
- Eller, C.B., Rowland, L., Mencuccini, M., Rosas, T., Williams, K., Harper, A., Medlyn, B.E., Wagner, Y., Klein, T., Teodoro, G.S., Oliveira, R.S., Matos, I.S., Rosado, B.H.P., Fuchs, K., Wohlfahrt, G., Montagnani, L., Meir, P., Sitch, S., Cox, P.M., 2020. Stomatal optimization based on xylem hydraulics (SOX) improves land surface model simulation of vegetation responses to climate. *New Phytol.* 226 (6), 1622–1637. <http://dx.doi.org/10.1111/nph.16419>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nph.16419>. Publisher: Blackwell Publishing Ltd.
- Feldman, A.F., Chulakadabba, A., Short Gianotti, D.J., Entekhabi, D., 2021. Landscape-scale plant water content and carbon flux behavior following moisture pulses: From dryland to mesic environments. *Water Resour. Res.* 57 (1), e2020WR027592. <http://dx.doi.org/10.1029/2020WR027592>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1029/2020WR027592>. Publisher: John Wiley & Sons, Ltd.
- Feng, X., Thompson, S.E., Woods, R., Porporato, A., 2019. Quantifying asynchronicity of precipitation and potential evapotranspiration in mediterranean climates. *Geophys. Res. Lett.* 46 (24), 14692–14701. <http://dx.doi.org/10.1029/2019GL085653>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1029/2019GL085653>. Publisher: John Wiley & Sons, Ltd.
- Fu, Z., Ciais, P., Feldman, A.F., Gentile, P., Makowski, D., Prentice, I.C., Stoy, P.C., Bastos, A., Wigneron, J.-P., 2022a. Critical soil moisture thresholds of plant water stress in terrestrial ecosystems. *Sci. Adv.* 8, 7827, URL: <https://www.science.org>.
- Fu, Z., Ciais, P., Makowski, D., Bastos, A., Stoy, P.C., Ibrom, A., Knohl, A., Migliavacca, M., Cuntz, M., Sigt, M., Peichl, M., Loustau, D., El-Madany, T.S., Buchmann, N., Gharun, M., Janssens, I., Markwitz, C., Grünwald, T., Rebmann, C., Mölder, M., Varlagin, A., Mammarella, I., Kolari, P., Bernhofer, C., Helias, M., Vincke, C., Pitacco, A., Cremonese, E., Foltynová, L., Wigneron, J.P., 2022b. Uncovering the critical soil moisture thresholds of plant water stress for European ecosystems. *Global Change Biol.* 28 (6), 2111–2123. <http://dx.doi.org/10.1111/GCB.16050>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.16050>. Publisher: John Wiley & Sons, Ltd.
- Fu, Z., Ciais, P., Prentice, I.C., Gentile, P., Makowski, D., Bastos, A., Luo, X., Green, J.K., Stoy, P.C., Yang, H., Hajima, T., 2022c. Atmospheric dryness reduces photosynthesis along a large range of soil water deficits. *Nat. Commun.* 13 (1), 1–10. <http://dx.doi.org/10.1038/s41467-022-28652-7>, URL: <https://www.nature.com/articles/s41467-022-28652-7>. Publisher: Nature Publishing Group.
- Guillaume, J.H.A., Jakeman, J.D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M.C., Jakeman, A.J., Keesman, K.J., Razavi, S., Stigter, J.D., 2019. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environ. Model. Softw.* 119, <http://dx.doi.org/10.1016/j.envsoft.2019.07.007>.
- Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowitz, G., Ye, M., 2012. Towards a comprehensive assessment of model structural adequacy. *Water Resour. Res.* 48 (8), 8301. <http://dx.doi.org/10.1029/2011WR011044>, URL: <https://onlinelibrary.wiley.com/doi/full/10.1029/2011WR011044>. Publisher: John Wiley & Sons, Ltd.
- Hari, P., Makela, A., Korpilahti, E., Holmberg, M., 1986. Optimal control of gas exchange. *Tree Physiol.* <http://dx.doi.org/10.1093/treephys/2.1-2.3.169>, ISBN: 0829-318X.
- Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* 7 (1), 109. <http://dx.doi.org/10.1038/s41597-020-0453-3>, URL: <https://www.nature.com/articles/s41597-020-0453-3>.
- Irvine, J., Law, B.E., Anthoni, P.M., Meinzer, F.C., 2002. Water limitations to carbon exchange in old-growth and young ponderosa pine stands. *Tree Physiol.* <http://dx.doi.org/10.1093/treephys/22.2-3.189>.
- Irvine, J., Law, B.E., Kurpius, M.R., Anthoni, P.M., Moore, D., Schwarz, P.A., 2004. Age-related changes in ecosystem structure and function and effects on water and carbon exchange in ponderosa pine. *Tree Physiol.* 24, 753–763. <http://dx.doi.org/10.1093/treephys/24.7.753>, URL: <https://academic.oup.com/treephys/article-abstract/24/7/753/1663720>.
- Irvine, J., Law, B.E., Martin, J.G., Vickers, D., 2008. Interannual variation in soil CO₂ efflux and the response of root respiration to climate and canopy gas exchange in mature ponderosa pine. *Global Change Biol.* 14 (12), 2848–2859. <http://dx.doi.org/10.1111/j.1365-2486.2008.01682.x>.
- Kannenberg, S.A., Guo, J.S., Novick, K.A., L. Anderegg, W.R., Feng, X., Kennedy, D., Konings, A.G., Martínez-Vilalta, J., Matheny, A.M., Steven Kannenberg, C.A., 2021. Opportunities, challenges and pitfalls in characterizing plant water-use strategies. *Funct. Ecol.* 36, 24–37. <http://dx.doi.org/10.1111/1365-2435.13945>, URL: <https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2435.13945>.
- Kennedy, D., Swenson, S., Oleson, K.W., Lawrence, D.M., Fisher, R., Lola da Costa, A.C., Gentile, P., 2019. Implementing plant hydraulics in the community land model, version 5. *J. Adv. Modelling Earth Syst.* 11 (2), 485–513. <http://dx.doi.org/10.1029/2018MS001500>, URL: <http://doi.wiley.com/10.1029/2018MS001500>. Publisher: John Wiley & Sons, Ltd.

- Kimm, H., Guan, K., Gentine, P., Wu, J., Bernacchi, C.J., Sulman, B.N., Griffis, T.J., Lin, C., 2020. Redefining droughts for the U.S. Corn Belt: The dominant role of atmospheric vapor pressure deficit over soil moisture in regulating stomatal behavior of Maize and Soybean. *Agricult. Forest Meteorol.* 287, <http://dx.doi.org/10.1016/j.agrformet.2020.107930>, Publisher: Elsevier B.V..
- Knauer, J., Zaehle, S., Medlyn, B.E., Reichstein, M., Williams, C.A., Migliavacca, M., De Kauwe, M.G., Werner, C., Keitel, C., Kolari, P., Limousin, J.-M., Linderon, M.-L., 2018. Towards physiologically meaningful water-use efficiency estimates from eddy covariance data. *Global Change Biol.* 24 (2), 694–710. <http://dx.doi.org/10.1111/gcb.13893>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.13893>. Publisher: John Wiley & Sons, Ltd (10.1111).
- Koster, R.D., Schubert, S.D., Suarez, M.J., 2009. Analyzing the concurrence of meteorological droughts and warm periods, with implications for the determination of evaporative regime. *J. Clim.* 22 (12), 3331–3341. <http://dx.doi.org/10.1175/2008JCLI2718.1>, URL: <https://journals.ametsoc.org/view/journals/clim/22/12/2008jcli2718.1.xml>. Publisher: American Meteorological Society.
- Lee, J.S., Choi, H.I., 2022. A rebalanced performance criterion for hydrological model calibration. *J. Hydrol.* 606, 127372. <http://dx.doi.org/10.1016/j.jhydrol.2021.127372>, Publisher: Elsevier.
- Li, X., Gentine, P., Lin, C., Zhou, S., Sun, Z., Zheng, Y., Liu, J., Zheng, C., 2019. A simple and objective method to partition evapotranspiration into transpiration and evaporation at eddy-covariance sites. *Agricult. Forest Meteorol.* 265, 171–182. <http://dx.doi.org/10.1016/j.agrformet.2018.11.017>, URL: <https://doi.org/10.1016/j.agrformet.2018.11.017>. Publisher: Elsevier.
- Lin, C., Gentine, P., Huang, Y., Guan, K., Kimm, H., Zhou, S., 2018. Diel ecosystem conductance response to vapor pressure deficit is suboptimal and independent of soil moisture. *Agricult. Forest Meteorol.* 250–251, 24–34. <http://dx.doi.org/10.1016/j.agrformet.2017.12.078>, URL: <https://www.sciencedirect.com/science/article/pii/S0168192317304884>. Publisher: Elsevier.
- Lin, Y.-S., Medlyn, B.E., Duursma, R.A., Prentice, I.C., Wang, H., Baig, S., Eamus, D., de Dios, V.R., Mitchell, P., Ellsworth, D.S., de Beek, M.O., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L.A., Nippert, J.B., Ocheltree, T.W., Tissue, D.T., Martin-StPaul, N.K., Rogers, A., Warren, J.M., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T.E., Barton, C.V.M., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S.A., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow, M.S.J., Drake, J.E., Freeman, M., Ghannoum, O., Hutley, L.B., Kelly, J.W., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola da Costa, A.C., Mikkelsen, T.N., Salinas, N., Sun, W., Wingate, L., 2015. Optimal stomatal behaviour around the world. *Nature Clim. Change* 5 (5), 459–464. <http://dx.doi.org/10.1038/nclimate2550>, URL: <http://www.nature.com/articles/nclimate2550>. Publisher: Nature Publishing Group.
- Liu, Y., Kumar, M., Katul, G.G., Feng, X., Konings, A.G., 2020. Plant hydraulics accentuates the effect of atmospheric moisture stress on transpiration. *Nature Clim. Change* 10, 691–695. <http://dx.doi.org/10.1038/s41558-020-0781-5>.
- Lloyd, J., Farquhar, G.D., Lloyd, J., Farquhar, G.D., 1994. 13C discrimination during CO₂ assimilation by the terrestrial biosphere. *Oecologia* 99 (3), 201–215. <http://dx.doi.org/10.1007/BF00627732>, URL: <https://link.springer.com/article/10.1007/BF00627732>. Publisher: Springer.
- Manzoni, S., Vico, G., Katul, G., Fay, P.A., Polley, W., Palmroth, S., Porporato, A., 2011. Optimizing stomatal conductance for maximum carbon gain under water stress: a meta-analysis across plant functional types and climates. *Funct. Ecol.* 25 (3), 456–467. <http://dx.doi.org/10.1111/j.1365-2435.2010.01822.x>, URL: <http://doi.wiley.com/10.1111/j.1365-2435.2010.01822.x>. Publisher: John Wiley & Sons, Ltd (10.1111).
- Matheny, A.M., Mirfenderesgi, G., Bohrer, G., 2017. Trait-based representation of hydrological functional properties of plants in weather and ecosystem models. *Plant Divers.* 39 (1), 1–12. <http://dx.doi.org/10.1016/J.PLD.2016.10.001>, URL: <https://www.sciencedirect.com/science/article/pii/S2468265916300154>. Publisher: Elsevier.
- Mauder, M., Foken, T., Cuxart, J., 2020. Surface-energy-balance closure over land: A review. *Bound.-Layer Meteorol.* 177 (2), 395–426. <http://dx.doi.org/10.1007/S10546-020-00529-6>, URL: <https://link.springer.com/article/10.1007/s10546-020-00529-6>.
- McPhail, C., Maier, H.R., Kwakkel, J.H., Giuliani, M., Castelletti, A., Westra, S., 2018. Earth's future robustness metrics: How are they calculated, when should they be used and why do they give different results? *Earth's Future* 6, 169–191. <http://dx.doi.org/10.1002/2017EF000649>, URL: <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017EF000649>.
- Medlyn, B.E., De Kauwe, M.G., Lin, Y.-S., Knauer, J., Duursma, R.A., Williams, C.A., Arneeth, A., Clement, R., Isaac, P., Limousin, J.-M., Linderson, M.-L., Meir, P., Martin-StPaul, N., Wingate, L., 2017. How do leaf and ecosystem measures of water-use efficiency compare? *New Phytol.* 216 (3), 758–770. <http://dx.doi.org/10.1111/nph.14626>, URL: <http://doi.wiley.com/10.1111/nph.14626>. Publisher: John Wiley & Sons, Ltd (10.1111).
- Medlyn, B.E., Duursma, R.A., Eamus, D., Ellsworth, D.S., Prentice, I.C., Barton, C.V.M., Crous, K.Y., De Angelis, P., Freeman, M., Wingate, L., 2011. Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biol.* 17 (6), 2134–2144. <http://dx.doi.org/10.1111/j.1365-2486.2010.02375.x>, URL: <http://doi.wiley.com/10.1111/j.1365-2486.2010.02375.x>. Publisher: Wiley/Blackwell (10.1111).
- Monteith, J.L., 1965. Evaporation and environment. *Symp. Soc. Exp. Biol.* 19, 205–234. <http://dx.doi.org/10.1613/jair.301>, arXiv:cs/9605103 ISBN: 0081-1386 (Print)/n0081-1386 (Linking).
- Muggeo, V.M.R., 2008. Segmented: an R package to fit regression models with broken-line relationships. *R News* 8 (1), 20–25, URL: <https://cran.r-project.org/doc/Rnews/>.
- Myneni, R., Knyazikhin, Y., Park, T., 2021. MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 global 500 m SIN grid V061 [Data set].
- Novick, K.A., Ficklin, D.L., Baldocchi, D., Davis, K.J., Ghezzehei, T.A., Konings, A.G., MacBean, N., Raoult, N., Scott, R.L., Shi, Y., Sulman, B.N., Wood, J.D., 2022. Confronting the water potential information gap. *Nat. Geosci.* 15 (3), 158–164. <http://dx.doi.org/10.1038/s41561-022-00909-2>, URL: <https://www.nature.com/articles/s41561-022-00909-2>. Publisher: Nature Publishing Group.
- 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., Phillips, R.P., 2016. The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Clim. Change* 6 (11), 1023–1027. <http://dx.doi.org/10.1038/nclimate3114>, Publisher: Nature Publishing Group.
- Novick, K.A., Konings, A.G., Gentine, P., 2019. Beyond soil water potential: An expanded view on isohydricity including land-atmosphere interactions and phenology. *Plant Cell Environ.* 42 (6), 1802–1815. <http://dx.doi.org/10.1111/pce.13517>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/pce.13517>. Publisher: Blackwell Publishing Ltd.
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M.A., Ardö, J., Arkebauer, T., Arndt, S.K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L.B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T.A., Blanken, P.D., Bohrer, G., Boike, J., Bolstad, P.V., Bonal, D., Bonnefond, J.-M., Bowling, D.R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S.P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T.R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B.D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P.S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K.J., Cinti, B.D., Grandcourt, A.D., Ligne, A.D., De Oliveira, R.C., Delpech, N., Desai, A.R., Di Bella, C.M., Tommasi, P.d., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., Elkhidir, H.A.M., Eugster, W., Ewen, C.M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Godeed, I., Goeckede, M., Goldstein, A.H., Gough, C.M., Goulden, M.L., Graf, A., Griebel, A., Gruening, C., Grünwald, T., Hammerle, A., Han, S., Han, X., Hansen, B.U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hörtnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janouš, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W.L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J.-M., Lion, M., Liska, A.J., Lohila, A., López-Ballesteros, A., López-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Lüers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H.A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J.H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A.M.S., Merbold, L., Meyer, W., Meyers, T., Miller, S.D., Minerbi, S., Moderow, U., Monson, R.K., Montagnani, L., Moore, C.E., Moors, E., Moreaux, V., Moureaux, C., Munger, J.W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosoet, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J.E., Ourcival, J.-M., Papuga, S.A., Parmentier, F.-J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R.P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S.M., Rambal, S., Rannik, U., Raz-Yaseef, N., Rebmann, C., Reed, D., Dios, V.R.D., Restrepo-Coupe, N., Reverter, B.R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S.R., Sánchez-Cañete, E.P., Sanchez-Mejia, Z.M., Schmid, H.P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R.L., Sedláč, P., Serrano-Ortiz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R.M., Sturtevant, C., Suyker, A., Tagessson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.-P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huisteden, K., Varlagin, A., Verfaillie, J., Versala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J.P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampieri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., Papale, D., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data* 7 (1), 1–27. <http://dx.doi.org/10.1038/s41597-020-0534-3>, URL: <https://www.nature.com/articles/s41597-020-0534-3>. Publisher: Nature Publishing Group.
- R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, URL: <https://www.r-project.org/>. Place: Vienna, Austria.
- Rowland, L., Ramírez-Valiente, J.-A., Hartley, I.P., Mencuccini, M., 2023. How woody plants adjust above- and below-ground traits in response to sustained drought. *New Phytol.* n/a (n/a), <http://dx.doi.org/10.1111/nph.19000>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nph.19000>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/nph.19000>.

- Ruddell, B.L., Drewry, D.T., Nearing, G.S., 2019. Information theory for model diagnostics: Structural error is indicated by trade-off between functional and predictive Performance. *Water Resour. Res.* 55, 6534–6554. <http://dx.doi.org/10.1029/2018WR023692>.
- Ruehr, N.K., Martin, J.G., Law, B.E., 2012. Effects of water availability on carbon and water exchange in a young ponderosa pine forest: Above- and belowground responses. *Agricult. Forest Meteorol.* 164, 136–148. <http://dx.doi.org/10.1016/j.agrformet.2012.05.015>.
- Sabot, M.E.B., De Kauwe, M.G., Pitman, A.J., Medlyn, B.E., Verhoef, A., Ukkola, A.M., Abramowitz, G., 2020. Plant profit maximization improves predictions of European forest responses to drought. *New Phytol.* 226 (6), 1638–1655. <http://dx.doi.org/10.1111/nph.16376>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nph.16376>. Publisher: Blackwell Publishing Ltd.
- Schwarz, P.A., Law, B.E., Williams, M., Irvine, J., Kurpius, M., Moore, D., 2004. Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems. *Glob. Biogeochem. Cycles* 18 (4), GB4007. <http://dx.doi.org/10.1029/2004GB002234>.
- Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Sci. Rev.* 99 (3–4), 125–161. <http://dx.doi.org/10.1016/j.earscirev.2010.02.004>.
- Sloan, B.P., 2023. Predictive and Explanatory Modeling of Ecosystem Response to Soil Water Stress (Ph.D. thesis). University of Minnesota -Twin Cities.
- Sloan, B.P., Thompson, S.E., Feng, X., 2021. Plant hydraulic transport controls transpiration sensitivity to soil water stress. *Hydrol. Earth Syst. Sci.* 25 (8), 4259–4274. <http://dx.doi.org/10.5194/HES-25-4259-2021>, Publisher: Copernicus GmbH.
- Starr, M.K., 1963. *Product Design and Decision Theory*. Prentice-Hall.
- Thomas, C.K., Law, B.E., Irvine, J., Martin, J.G., Pettijohn, J.C., Davis, K.J., 2009. Seasonal hydrology explains interannual and seasonal variation in carbon and water exchange in a semiarid mature ponderosa pine forest in central Oregon. *J. Geophys. Res.: Biogeosci.* 114 (4), <http://dx.doi.org/10.1029/2009JG001010>, Publisher: Blackwell Publishing Ltd.
- Wang, H., Yan, S., Ciais, P., Wigneron, J.-P., Liu, L., Li, Y., Fu, Z., Ma, H., Liang, Z., Wei, F., Wang, Y., Li, S., 2022. Exploring complex water stress-gross primary production relationships: Impact of climatic drivers, main effects, and interactive effects. *Glob. Change Biol.* 28, 4110–4123. <http://dx.doi.org/10.1111/gcb.16201>.
- Wolf, A., Anderegg, W.R.L., Pacala, S.W., 2016. Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proc. Natl. Acad. Sci. USA* 113 (46), E7222–E7230. <http://dx.doi.org/10.1073/pnas.1615144113>, URL: <http://www.ncbi.nlm.nih.gov/pubmed/27799540>. Publisher: National Academy of Sciences.
- Zhou, S., Duursma, R.A., Medlyn, B.E., Kelly, J.W., Prentice, I.C., 2013. How should we model plant responses to drought? An analysis of stomatal and non-stomatal responses to water stress. *Agricult. Forest Meteorol.* 182–183, 204–214. <http://dx.doi.org/10.1016/J.AGRFORMET.2013.05.009>, URL: <https://www.sciencedirect.com/science/article/pii/S0168192313001263>. Publisher: Elsevier.
- Zhou, S., Medlyn, B., Sabaté, S., Sperlich, D., Prentice, I.C., Whitehead, D., 2014. Short-term water stress impacts on stomatal, mesophyll and biochemical limitations to photosynthesis differ consistently among tree species from contrasting climates. *Tree Physiol.* 34 (10), 1035–1046. <http://dx.doi.org/10.1093/treephys/tpu072>, URL: <https://academic.oup.com/treephys/article-lookup/doi/10.1093/treephys/tpu072>. Publisher: Oxford University Press.
- Zhou, S., Yu, B., Huang, Y., Wang, G., 2015. Daily underlying water use efficiency for AmeriFlux sites. *J. Geophys. Res.: Biogeosci.* 120 (5), 887–902. <http://dx.doi.org/10.1002/2015JG002947>, URL: <http://public.ornl.gov/>. Publisher: Blackwell Publishing Ltd.