BES-manuscript

1 Introduction

Soils contain the largest reservoir of terrestrial carbon (Jobbágy & Jackson, 2000). Accumulation of soil organic matter can be influenced by biotic factors from above-ground plant inputs (Jackson et al., 2017), which in turn are influenced through environmental factors such as growing season length, temperature, and moisture (Desai et al., 2022). These biotic factors provide natural climate-based solutions to offsetting any net increases of the terrestrial carbon cycle from greenhouse gas emissions (Friedlingstein et al., 2023).

Ecological networks such as the National Ecological Observatory Network (NEON) and others (e.g. FLUXNET or the Integrated Carbon Observation System) present a significant advancement in the nearly continuous observation of biogeochemical processes at the continental scale. Notably, at 47 core terrestrial sites across the continental United States NEON provides half-hourly measurements of soil carbon content, soil CO_2 concentration, temperature, and moisture at different vertical depths. In turn, FLUXNET provides measurements of the cumulative sum of all ecosystem carbon fluxes in an airshed using the eddy covariance technique (Baldocchi, 2014). Soil observations from NEON data provided by NEON is on the same timescale and standardized with eddy covariance measurements from FLUXNET. When combined together NEON and FLUXNET data reconciling differences between model-derived or data-estimated components (Jian et al., 2022; Luo et al., 2011; Phillips et al., 2017; J. Shao et al., 2015; P. Shao et al., 2013; Sihi et al., 2016).

Beyond observations of soil CO_2 concentrations or other environmental variables such as moisture or temperature, a key quantity is soil carbon fluxes (Bond-Lamberty et al., 2024). A soil carbon flux to the atmosphere (F_S, units μmol m^{-2} s^{-1}) represents the aggregate process of transfer of soil CO_2 to the atmosphere from physical and biological processes (e.g. diffusion and microbial and autotrophic respiration). Soil carbon fluxes can be coupled with empirical or process models of soil carbon. Soil carbon fluxes can be assumed to represent soil carbon respiration from autotrophic or heterotrophic sources (Davidson et al., 2006), typically assumed to be static across the soil biome and modeled with a exponential Q_{10} paradigm (Bond-Lamberty et al., 2004; Chen & Tian, 2005; Hamdi et al., 2013).

Measurement of F_S is done through directly with soil chambers (e.g. LICOR or XXXX) or from soil measurements at different depths in the soil. In the latter case, the flux-gradient method is an approach derives from the mass conservation from the vertical soil depth z at steady state by applying Fick's law of diffusion. A simplifying assumption for the flux-gradient method is that there no mass transfer in the other spatial dimensions x and y (Maier & Schack-Kirchner, 2014). The diffusivity profile across the soil depth is a function of soil temperature, soil volumetric water content, atmospheric air pressure, and soil bulk density (Millington & Shearer, 1971; Moldrup et al., 1999).

A growing number of databases such as the Soil Respiration Database (SRDB) or Continuous Soil Respiration Database (COSORE) add to the growing network of observations of soil fluxes (Bond-Lamberty, 2018; Bond-Lamberty et al., 2020; Bond-Lamberty & Thomson, 2010; Jian et al., 2021; Jiang et al., 2024). Currently NEON provides all measurements to calculate F_S from Fick's law, but it was descoped from the initial network launch (Berenbaum et al., 2015). Deriving data-based measurements of F_S is a high priority.

This study describes efforts to derive a standardized estimate of F_S at all terrestrial NEON sites. Derived values of F_S are then validated from a subset of field observations. Key objectives of this study are to:

- apply the flux-gradient method to measurement to current NEON sites
- benchmark produced soil carbon fluxes to other ancillary measurements (e.g. SRDB, measurements of soil respiration)
- identify sources of error for future work.

2 Materials and Methods

2.1 Study sites

We selected six terrestrial NEON sites for analysis. These sites span a range of environmental gradients and terrestrial domains for analysis (Table 1). Over the course of two field campaigns in 2022 and 2024 we conducted weekly visits at each site through selecting a specific in the soil sampling array, installing a temporary soil collar, and doing direct flux measurements. These data were then compared for analysis later.

Table 1: Listing of NEON sites studied for field work and analysis.

Site	Location	Ecosystem type	Mean annual temperature (sampling)	Mean annual precipitation (sampling)
Santa Rita Experimental Range	31.91068, -110.83549	Shrubland	19.3°C	346.2 mm
San Joaquin Experimental Range	37.10878, -119.73228	Oak woodland	16.4°C	539.62 mm
Wind River Experimental Forest	45.82049, -121.95191	Evergreen forest	9.2°C	2225 mm
Chase Lake National Wildlife Refuge	47.1282, -99.241334	Restored prairie grassland	4.9°C	495 mm
Konza Prairie Biological Station	39.100774, -96.563075	Tallgrass Prairie	12.4°C	870 mm
University of Notre Dame Environmental Research Center	46.23391, -89.537254	Deciduous forest	4.3°C	802 mm

IDEA: Could have a sparkline plot of temperature / precip over the course of the time spent at each site, although admittedly we didn't see THAT much variation when sampling.

2.2 Field methods

In order to acquire field data to validate model predictions of flux, we conducted field measurement campaigns at the siz core terrestrial NEON sites listed above. SJER, SRER, and WREF were visited during May and June of 2022, and WOOD, KONZ, and UNDE during May and June of 2024. We spent a week at each site, taking daily measurements of flux on an hourly or half-hourly interval after letting soil collar(s) equilibrate for approximately 24 hours.

2.2.1 Soil collar placement

Either one (2022 sampling campaign) or two (2024 sampling campaign) PVC soil collars (FIXME: diameter) were installed in close proximity to the permanent NEON soil sensors

at each site. The soil plot where measurements was taken was chosen at each site in consultation with NEON staff to maximize likelihood of quality soil sensor measurements during the duration of the IRGA measurements at each site.

IDEA: Add graphic of soil plot layout and placement of soil collar(s) – could make diagram in OmniGraffle?

2.2.2 Infrared gas analyzer measurements of soil CO_2 flux

During the summer 2022 field campaign, a LI-COR 6800 with soil flux chamber attachment was used to measure soil fluxes for 8 hours each day on an hourly interval. During the summer 2024 field campaign, the LI-6800 measurements were taken on a half-hourly interval and were paired with an automated soil flux chamber setup (FIXME multiplexer, IRGA, chamber model numbers) that made automated measurements on a half-hourly interval 24 hours a day while we were on site. Each instrument was paired with a soil temperature and moisture probe (FIXME: Stevens model #) that was used to make soil temperature and moisture measurements concurrent with the CO_2 flux measurements.

Dead bands, measurement duration, instrument self-testing.

2.2.3 Post-collection processing of data

LI-COR SoilFluxPro software to assess dead band and measurement duration.

2.3 neonSoilFlux R package

We developed an R package (`neonSoilFlux`; https://CRAN.R-project.org/package=neonSoilFlux) to compute half-hourly soil carbon fluxes and uncertainties from NEON data. The objective of the `neonSoilFlux` package is a unified workflow soil data acquisition and analysis that supplements existing data acquisition software through the `neonUtilities` R package (https://CRAN.R-project.org/package=neonUtilities). Figure 1 outlines the basic workflow of the package.

At a given NEON observation there are five different replicate soil sensor arrays, with each sampling at five different soil depths (Figure 2). The `neonSoilFlux` package acquires measured soil water content (National Ecological Observatory Network (NEON), 2024e), soil CO_2 concentration (National Ecological Observatory Network (NEON), 2024b), barometric pressure (National Ecological Observatory Network (NEON), 2024a), soil temperature (National Ecological Observatory Network (NEON), 2024d), and soil properties (e.g. bulk density) (National Ecological Observatory Network (NEON), 2024c). The static soil properties are periodically collected and assumed to be constant for the monthly observation period.

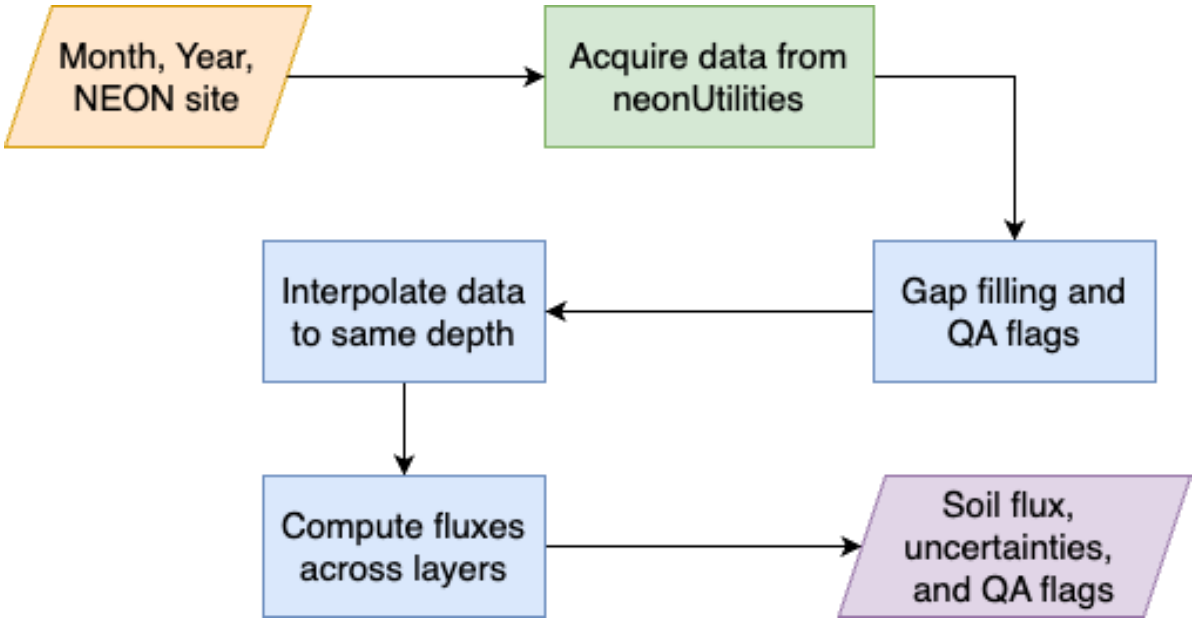

Figure 1: Diagram of `neonSoilFlux` R package. For a given month and NEON site, the package acquires all relevant data to compute F_S using the `neonUtilities` R package. Data are gap-filled according to reported QA flags and interpolated to the same measurement depth before computing the soil flux, uncertainties, and final QA flags.

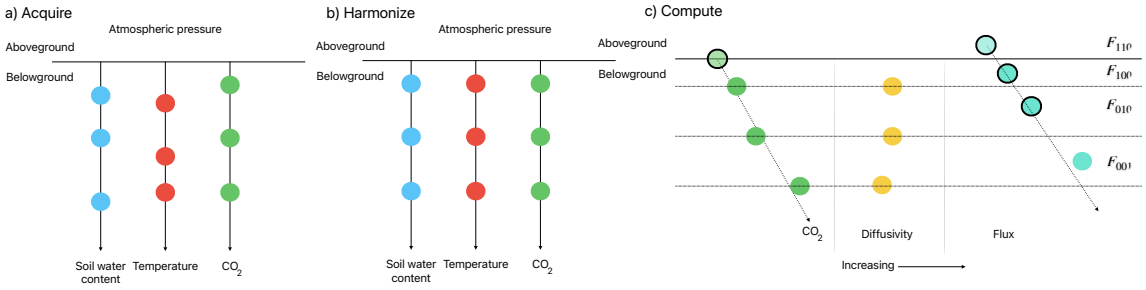

Figure 2: Model diagram for data workflow for the neonSoilFlux R package. a) Acquire: Data are obtained from given NEON location and horizontal sensor location, which includes soil water content, soil temperature, CO_2 concentration, and atmospheric pressure. All data are screened for quality assurance, with gap-filling of missing data reported. b) Any belowground data are then harmonized to the same depth as CO_2 concentrations using linear regression. c) The flux across a given depth is computed via Fick's law, denoted with F_{ijk}, where i, j, or k are either 0 or 1 denoting the layer the flux is computed across (i = closest to surface, k = deepest). The surface flux is all possible combinations of F_{ijk} extrapolating the flux measurements to the surface, so F_{110} is the surface flux intercept linearly extrapolating the measurements F_{010} and F_{100}.

The workflow to computing a value of F_S with the `neonSoilFlux` consists of three primary steps. First, NEON data are acquired for a given site and month via the `neonUtilities` R package (yellow parallelogram and green rectangle in Figure 1 and Panel a in Figure 2). Acquired environmental data can be exported to a comma separated value file for additional analysis. Quality assurance (QA) flags with an observation are reported as an indicator variable.

The next step is harmonizing the data to compute soil fluxes across soil layers. This step consists of three different actions (blue rectangles in Figure 1 and Panel b in Figure 2). If a given observation did is reported as not passing a quality assurance check we applied a gap filling method to replace that measurement with its monthly mean at that same depth (Section 2.3.1). Belowground measurements of soil water and soil temperature are then interpolated to the same depth as soil CO_2 measurements. The diffusivity (Section 2.3.2) and soil flux across different soil layers (Section 2.3.3) are then computed.

The final step is computing a surface soil flux through extrapolation to the surface (purple parallelogram in Figure 1 and Panel c in Figure 2). Uncertainty on a soil flux measurement is computed through quadrature. An aggregate QA flag for each environmental measurement is also reported, representing if any gap-filled measurements were used in the computation of a soil flux. Within the soil flux-gradient method, several different approaches can be used to derive a surface flux (Maier & Schack-Kirchner, 2014); the `neonSoilFlux` package reports

eight different possible values of soil surface flux (Section 2.3.3).

2.3.1 Gap-filling routine

NEON reports QA flags as a binary value for a given measurement and half-hourly time. We replaced any flagged measurements at a location's spatial depth z with a bootstrapped sample of the monthly mean for all un-flagged measurements for that month. hese measurements are represented by the vector $\mathbf{m}$, standard errors σ, and the 95% confidence interval (the so-called expanded uncertainty, Farrance & Frenkel (2012)) ϵ. All of these vectors have length M. We have that $\vec{\sigma}_i \leq \vec{\epsilon}_i$. We define the bias as $\mathbf{b} = \sqrt{\epsilon^2 - \sigma^2}$.

We generate a vector of bootstrap samples of the distribution of the monthly mean $\overline{m}$ and monthly standard error $\overline{\sigma}$ the following ways:

1. Randomly sample from the uncertainty and bias independently: σ_j and the bias $\mathbf{b}_k$ (not necessarily the same sample)
2. Generate a vector $\mathbf{n}$ of length N, where $\mathbf{n}_i$ is a random sample from a normal distribution with mean m_i and standard deviation σ_j. Since $M < N$, values from $\mathbf{m}$ will be reused.
3. With these N random samples, $\overline{y}_i = \overline{\vec{x}} + \vec{b}_k$ and s_i is the sample standard deviation of $\vec{x}$. We expect that $s_i \approx \vec{\sigma}_j$.
4. The reported monthly mean and standard deviation are then computed $\overline{\overline{y}}$ and $\overline{s}$. Measurements and uncertainties that did not pass the QA check are then substituted with $\overline{\overline{y}}$ and $\overline{s}$.

2.3.2 Diffusivity computation

Soil diffusivity D_a at a given measurement depth is the product of the diffusivity in free air $D_{a,0}$ ($m^2\ s^{-1}$) and the tortuosity ξ (no units) (Millington & Shearer, 1971). Surface barometric pressure (kPa) (National Ecological Observatory Network (NEON), 2024a), soil temperature at depth (National Ecological Observatory Network (NEON), 2024d), soil water content (National Ecological Observatory Network (NEON), 2024e), and soil physical properties (National Ecological Observatory Network (NEON), 2024c). Soil physical properties are surveyed once at each site, whereas the other measurements are provided on a half-hourly basis.

We compute $D_{a,0}$ with Equation 1:

$$D_{a,0} = 0.0000147 \cdot \left(\frac{T_i + 273.15}{293.15}\right)^{1.75} \cdot \left(\frac{P}{101.3}\right) \tag{1}$$

where T_i is soil temperature (°C) at depth i (National Ecological Observatory Network (NEON), 2024d) and P surface barometric pressure (kPa) (National Ecological Observatory

Network (NEON), 2024a). At that soil depth, the tortuosity ξ is defined by Equation 2 (Millington & Shearer, 1971):

$$\xi = \frac{(\phi - SWC_i)^{10/3}}{\phi^2} \tag{2}$$

In Equation 2, SWC is the soil water content at depth i (National Ecological Observatory Network (NEON), 2024e) and ϕ is the porosity (Equation 3), which in turn is a function of soil physical properties (National Ecological Observatory Network (NEON), 2024c).

The tortuosity ξ is computed from soil porosity ϕ. :

$$\phi = \left(1 - \frac{\rho_s}{\rho_m}\right)(1 - f_V) \tag{3}$$

In Equation 3, ρ_m is the particle density of mineral soil (2.65 g cm^{-3}), ρ_s the soil bulk density (g cm^{-3}) excluding coarse fragments greater than 2 mm (National Ecological Observatory Network (NEON), 2024c). The term f_V is a site-specific value that accounts for the proportion of soil fragments between 2-20 mm. Soil fragments greater than 20 mm were not estimated due to limitations in the amount of soil that can be analyzed (National Ecological Observatory Network (NEON), 2024c). We assume there are no pores within rocks. Values of ρ_s, ρ_m are assumed to be constant across the soil profile and the same at each site sampling location.

2.3.3 Soil flux computation

We applied Fick's law (Equation 4) to compute the soil flux F_{ij} (μmol m^{-2} s^{-1}) across two adjacent soil depths i and j:

$$F_{ij} = -D_a \frac{dC}{dz} \tag{4}$$

where D_a is the diffusivity (m^2 s^{-1}) and $\frac{dC}{dz}$ is the gradient of CO_2 molar concentration (μmol m^{-3}, so the gradient has units of μmol m^{-3} m^{-1}). The diffusivity (described below) is a function of soil temperature, soil water content, and soil physical properties. The soil surface flux is theoretically defined by applying Equation 4 to measurements collected at the soil surface and directly below the surface. Measurements of soil temperature, soil water content, and soil CO_2 molar concentration across the soil profile allow for application of Equation **??** across different soil depths. The flux gradient method approximates the soil surface flux either by (1) extrapolation of Equation 4 across sub-surface measurement depths to the surface, typically assuming soil flux is a linear function of depth (Maier & Schack-Kirchner, 2014) or (2) linear extrapolation of D_a to the surface and from direct calculation of $\frac{dC}{dz}$ from the CO_2 profile. All these approaches are pThe `neonSoilFlux` package provides several different methods to compute F_s for the end-user to compare.

2.3.4 Reporting of surface fluxes

A surface flux estimate is derived from Fick's Law (Equation 4), which is the product of a diffusivity and a CO_2 concentration gradient (Maier & Schack-Kirchner, 2014). The `neonSoilFlux` package provides eight different surface flux estimates, which represent different considerations of how Fick's Law is applied. First, we apply simple linear regression to both CO_2 and D_a at the three different measurement depths. Next, the slope and intercept (and uncertainty by quadrature) from these regressions are used to compute a suite of eight different surface flux estimates (denoted by F_{ijk}):

- F_{000} is a surface flux estimate using the intercept of the linear regression of D_a and the slope from linear regression of CO_2 (which represents $\frac{dC}{dz}$ in Fick's Law). Tang et al. (2003) used this approach to compute fluxes in an oak-grass savannah.
- F_{010}, F_{001} are fluxes across the two most shallow layers and two deepest layers respectively. The diffusivity used in Fick's Law is always at the deeper measurement layer. When used as a surface flux estimate we assume CO_2 remains constant above this flux depth.
- F_{100} is a flux estimate where the gradient $\frac{dC}{dz}$ is estimated using the intercept from linear regression of CO_2 and the top measurement depth for CO_2. The diffusivity used in Fick's Law is always at the first measurement layer. de Jong & Schappert (1972) applied this approach in a Canadian prairie.
- F_{110}, F_{101}, F_{011} are a surface flux estimates using linear extrapolation between F_{100} and F_{010}; F_{100} and F_{001}; or F_010 and F_{001} respectively. Hirano et al. (2003) and Tang et al. (2005) used an approach similar to F_{101} in a temperate deciduous broadleaf forest and ponderosa pine forest respectively.
- F_{111} is a surface flux estimate using linear extrapolation between F_{100}, F_{010}, and F_{001}.

Uncertainty in all F_{ijk} is computed through quadrature.

2.4 Post processing evaluation

Following collection of field measurements from the LICOR and calculation of the soil fluxes from `neonSoilFlux` package, we compared measured soil fluxes (from the LICOR instruments) to a given soil flux calculation during the same half hour. At each site and flux compuation method we also computed the R^2 value for the half-hourly and measured LICOR values and the RMSE.

We evaluate the efficacy of results from the flux-gradient method in two ways. First, we calculated the signal to noise ratio (SNR), defined as the ratio of the reported flux to its uncertainty (F_{ijk}/σ_{ijk}).

We evaluated if the measured field fluxes were within the calculated uncertainty from the flux-gradient method using the various approaches outlined above. We observed that the calculated

quadrature uncertainty in many cases can be much larger than the reported measurement (as shown through the signal to noise ratio, SNR $= F_{ijk}/\sigma_{ijk}$. We evaluated $|F_S - F_{ijk}| < (1-\epsilon)\sigma_{ijk}$, where F_S is a measured field soil flux from the LICOR 6800 (the LICOR 8250 was used at only three sites) and F_{ijk} is a computed flux method from the flux-gradient, and σ_{ijk} is the reported uncertainty for the flux method. The parameter ϵ was an uncertainty reduction factor to evaluate how sensitive the results were given, measured by the proportion of field measurements contained in that range.

2.5 Diffusivity back calculations

We also computed the diffusivity at each of the different sites from the derived flux values and the reported gradient in CO2 at each of the surface layers (Figure XXXX) through dividing $F_{field}/dC/dz$ These values were then compared to reported diffusivity from the flux computation method.

Finally, the bulk density at a given site was then also computed through back calculation.

3 Results

Figure Figure 3 reports out the measured fluxes from the LICOR 6800 and 8250 and computed fluxes and uncertainty at each measurement site. Results are reported in local time. Positive values of the flux indicate that there is a flux moving towards the surface (FIX WRITING) For ease of clarity the fluxes at F_{111} and F_{000} are only shown in the top row (surface), followed by the fluxes at individual separate layer (F_{100}, F_{010}, F_{001}). Overall, with the exception of WREF and SRER (discussed later) the computed fluxes were on the same order of magnitude and timing as the measured field fluxes.

Figure Figure 4 reports which environmental measurements at each site relied on gap-filled measurements (left plot) and in the aggregate, the total number of gap-filled measurements used for each half-hourly flux computation. the largest contribution to gap-filled data was soil water follwed by soil temperature.

Figure: flux results at the different levels (000,111,001,010,100) Diffusivity at the different levels for comparison (also include derived diffusivity?) Stats at the different levels (with the lags)

LEAD IN TABLE HERE These results are reported in Figure 6.

Calculated diffusivity from both the `neonSoilFlux` method and back calculated from the LICOR are reported in Figure Figure 7. SAY MORE

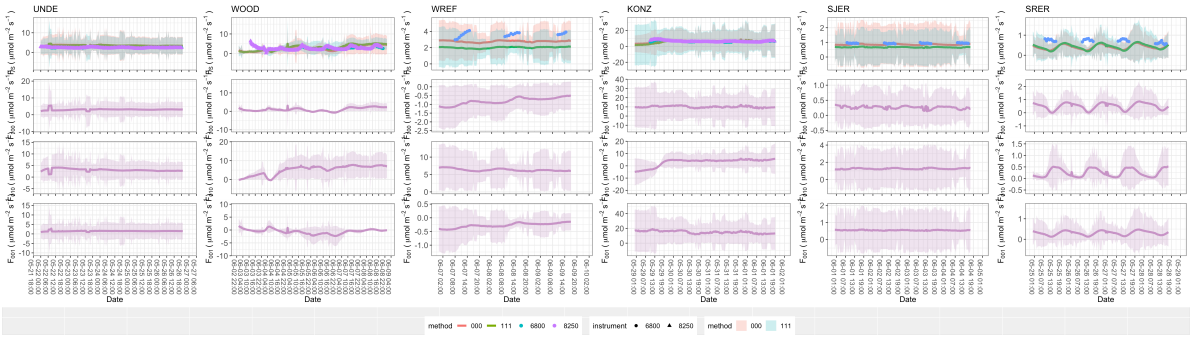

Figure 3: Results for different flux computations, organized by site (in increasing temperature show at each site) across different measurement levels. F_{000} comes from the diffusivity extrapolation and F_{111} extrapolation across the surface. Field measurements are shown at the top of each plot. The computed flux values are shown with reported uncertainty as well.

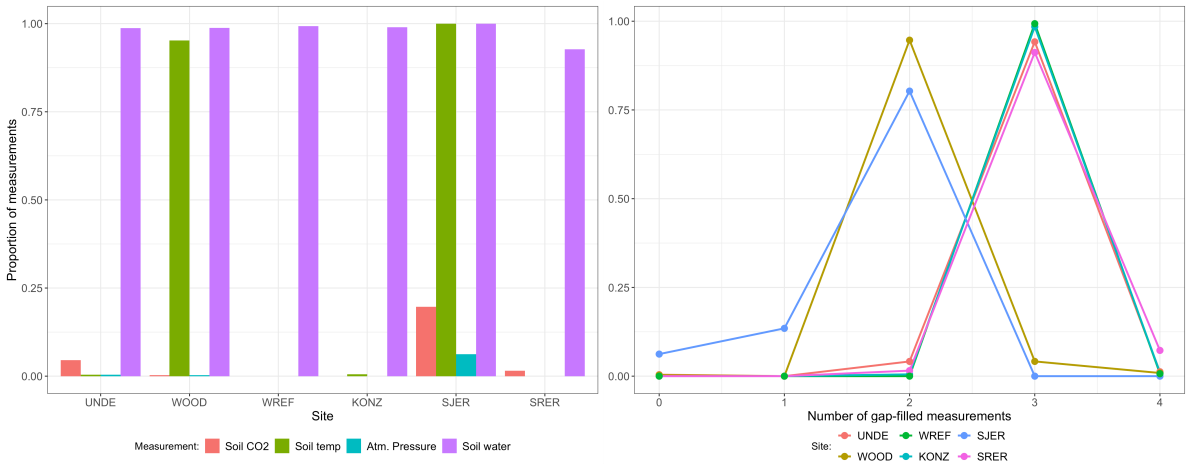

Figure 4: Left panel: Proportion of gap-filled environmental observations at each site. Right panel: proportion of measurements that used gap-filled environmental data (0 - 4) at each site. BLAH

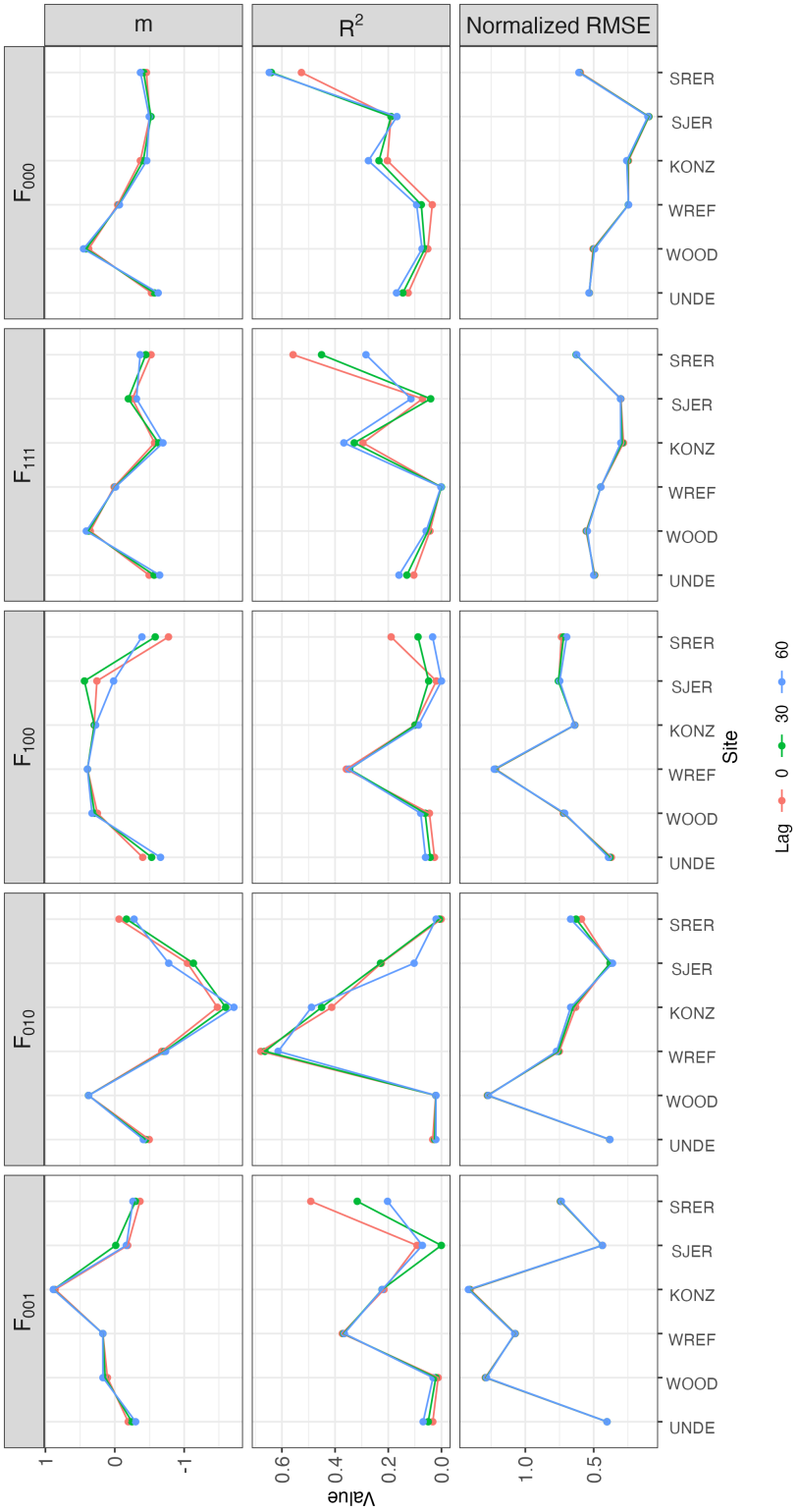
m
R2
Normalized RMSE
F000
F111
F100
F010
F001
SRER
SJER
KONZ
WREF
WOOD
UNDE
Site
Lag 0 30 60
Value
1
0
-1
0.6
0.4
0.2
0.0
1.0
0.5

Figure 5

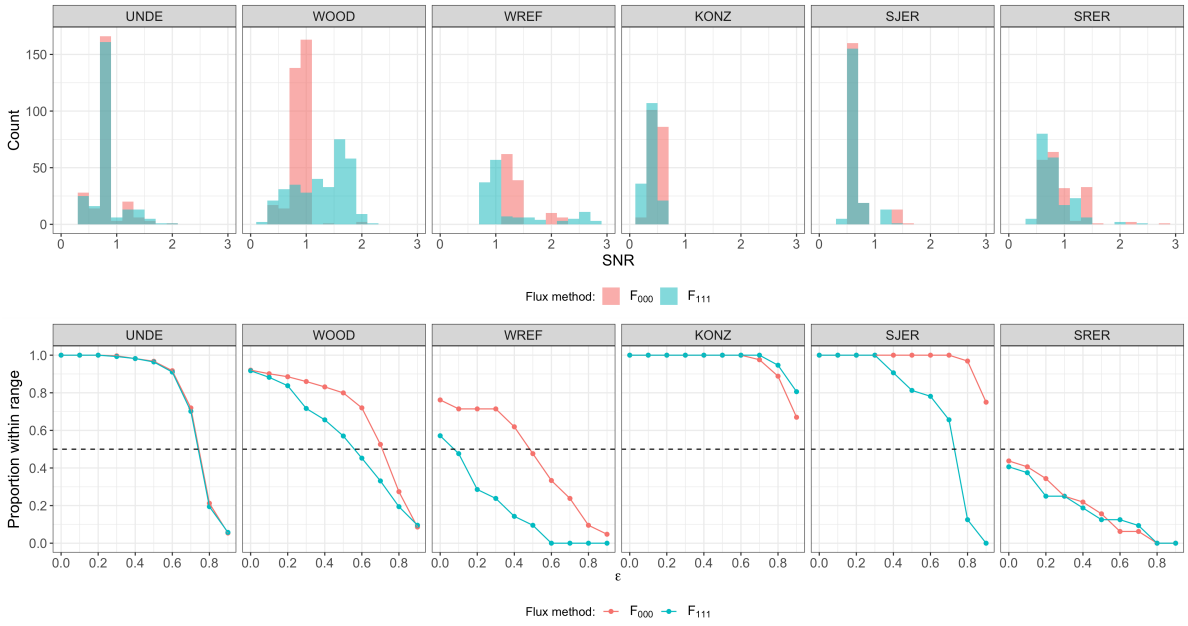

Figure 6: Top panel: distribution of SNR values across each of the different sites for the F_{000} and F_{111} flux gradient calculations. Bottom panel: Computation of uncertainty reduction to evaluate. As ϵ increases this indicates that the uncertainty estimate reduces, making it harder to be within the range. BLAH

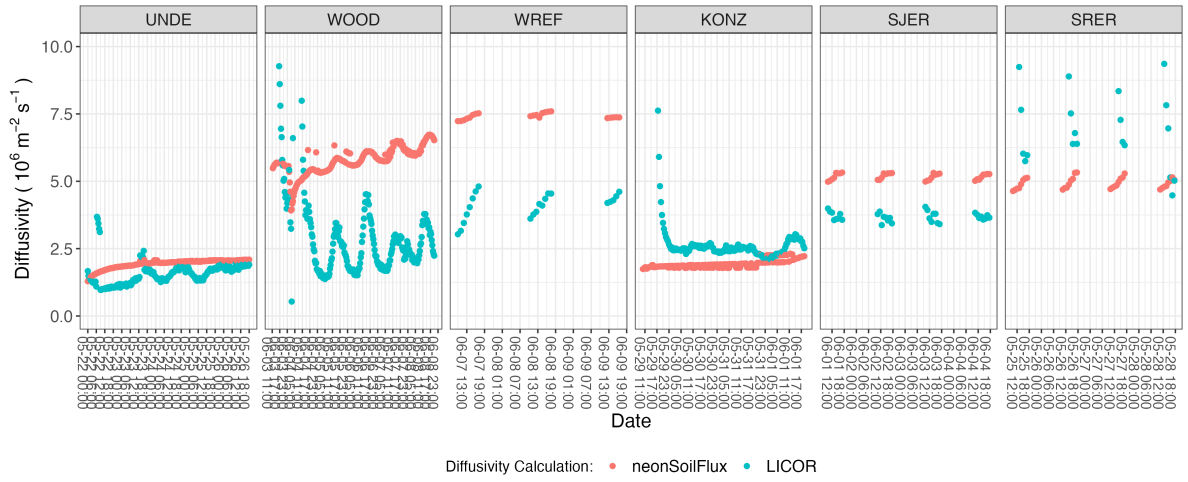

Figure 7: Calculation of diffusivity at each of the sites SAY MORE

4 Discussion

This study presents a unified data science workflow to efficiently process automated measurements of belowground soil CO_2 concentrations, water, and temperature to infer estimates of soil effluxes through application of Fick's Law. Derived results from the `neonSoilFlux` package have patterns that are consistent, and comparable, to those directly measured to the field (Figure XXX). The advantage to the `neonSoilFlux` package is the calculation of fluxes across different measurement depths, allowing for additional site-specific customization. Here application of the flux-gradient method provides a baseline estimate of soil fluxes that could be complemented through additional field measurements (e.g. LICOR).

The largest source of uncertainty to improve reliability of the flux estimate is to prevent the usage of gap-filled data, especially with soil water (Figure XXX). Across the observed half-hourly periods for field measurements, the percentage of half-hourly periods where all four environmental measurements were available spanned from 0% (SJER) to 7% (SRER), but three sites (WREF, SRER, UNDE) had 95% of half-hourly intervals with just one gap-filled measurement.Where appropriate we have replaced measurements flagged to the protocol with a monthly gap-filled value. Further extensions of the gapfilling method could use more sophisticated gap-filling routines, similar to what is used for net ecosystem carbon exchange (Falge et al., 2001; Liu et al., 2023; Mariethoz et al., 2015; Moffat et al., 2007; Zhang et al., 2023).

The `neonSoilFlux` package provides eight different approaches of values for a soil flux. We believe these approaches reflect a variety of site-specific determination and assumptions used to generate a soil flux measurement (Maier & Schack-Kirchner, 2014), with the choice of method having a determinative approach on reported values. Reported results could further be distilled down using ensemble averaging averaging approaches (Elshall et al., 2018; Raftery et al., 2005).

Figures XXX suggests that the provided uncertainty from `neonSoilFlux` is an overestimate compared to what is actually computed. When $\epsilon = 0$ in Figure Figure 6, that means we are just using the reported uncertainty from neonSoilFlux. Looking at that (epsilon = 0) shows field measurements UNDE, KONZ, SJER are 100% within the reported intervals from neonSoilFlux. But those sites tend to have a SNR < 1, so the uncertainty is pretty noisy. For UNDE, we could even reduce the uncertainty by a factor of 75% (epsilon = 0.75), more than half of the field measurements will still be within the reported intervals. For KONZ, we are still within 70% of the reported intervals when uncertainty is reduced by 90%. To me, that suggests that while the reported accuracy (as compared to field measurements), we do have higher precision.

Diffusivity taken from the megapit. Big microsite and temporal differences, soil shrink / swell. How do we think about what that means for using this approach to generalize across a NEON site over continuous time.

- Back calculation of bulk density to NEON sites.

5 Conclusions

6 Acknowledgments

7 Conflict of Interest Statements

Author Contributions

Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere - the state and future of the eddy covariance method. *Global Change Biology*, *20*(12), 3600–3609. https://doi.org/10.1111/gcb.12649

Berenbaum, M. R., Carpenter, S. R., Hampton, S. E., Running, S. W., & Stanzione, D. C. (2015). *Report from the NSF BIO Advisory Committee Subcommittee on NEON Scope Impacts*.

Bond-Lamberty, B. (2018). New Techniques and Data for Understanding the Global Soil Respiration Flux. *Earth's Future*, *6*(9), 1176–1180. https://doi.org/10.1029/2018EF000866

Bond-Lamberty, B., Ballantyne, A., Berryman, E., Fluet-Chouinard, E., Jian, J., Morris, K. A., Rey, A., & Vargas, R. (2024). Twenty Years of Progress, Challenges, and Opportunities in Measuring and Understanding Soil Respiration. *Journal of Geophysical Research: Biogeosciences*, *129*(2), e2023JG007637. https://doi.org/10.1029/2023JG007637

Bond-Lamberty, B., Christianson, D. S., Malhotra, A., Pennington, S. C., Sihi, D., AghaKouchak, A., Anjileli, H., Altaf Arain, M., Armesto, J. J., Ashraf, S., Ataka, M., Baldocchi, D., Andrew Black, T., Buchmann, N., Carbone, M. S., Chang, S.-C., Crill, P., Curtis, P. S., Davidson, E. A., … Zou, J. (2020). COSORE: A community database for continuous soil respiration and other soil-atmosphere greenhouse gas flux data. *Global Change Biology*, *26*(12), 7268–7283. https://doi.org/10.1111/gcb.15353

Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Biogeosciences*, *7*(6), 1915–1926. https://doi.org/10.5194/bg-7-1915-2010

Bond-Lamberty, B., Wang, C., & Gower, S. T. (2004). A global relationship between the heterotrophic and autotrophic components of soil respiration? *Global Change Biology*, *10*(10), 1756–1766. https://doi.org/10.1111/j.1365-2486.2004.00816.x

Chen, H., & Tian, H.-Q. (2005). Does a General Temperature-Dependent Q10 Model of Soil Respiration Exist at Biome and Global Scale? *Journal of Integrative Plant Biology*, *47*(11), 1288–1302. https://doi.org/10.1111/j.1744-7909.2005.00211.x

Davidson, E. A., Janssens, I. A., & Luo, Y. (2006). On the variability of respiration in terrestrial ecosystems: Moving beyond Q10. *Global Change Biology*, *12*, 154–164. https://doi.org/10.1111/j.1365-2486.2005.01065.x

de Jong, E., & Schappert, H. J. V. (1972). Calculation of Soil Respiration and Activity from CO2 Profiles in the Soil. *Soil Science*, *113*(5), 328–333.

Desai, A. R., Murphy, B. A., Wiesner, S., Thom, J., Butterworth, B. J., Koupaei-Abyazani, N., Muttaqin, A., Paleri, S., Talib, A., Turner, J., Mineau, J., Merrelli, A., Stoy, P., & Davis, K. (2022). Drivers of Decadal Carbon Fluxes Across Temperate Ecosystems. *Journal of Geophysical Research: Biogeosciences*, *127*(12), e2022JG007014. https://doi.org/10.1029/2022JG007014

Elshall, A. S., Ye, M., Pei, Y., Zhang, F., Niu, G.-Y., & Barron-Gafford, G. A. (2018). Relative model score: A scoring rule for evaluating ensemble simulations with application to microbial soil respiration modeling. *Stochastic Environmental Research and Risk Assessment*, *32*(10), 2809–2819. https://doi.org/10.1007/s00477-018-1592-3

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., … Wofsy, S. (2001). Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural and Forest Meteorology*, *107*(1), 43–69. https://doi.org/10.1016/S0168-1923(00)00225-2

Farrance, I., & Frenkel, R. (2012). Uncertainty of Measurement: A Review of the Rules for Calculating Uncertainty Components through Functional Relationships. *The Clinical Biochemist Reviews*, *33*(2), 49–75.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., … Zheng, B. (2023). Global Carbon Budget 2023. *Earth System Science Data*, *15*(12), 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Hamdi, S., Moyano, F., Sall, S., Bernoux, M., & Chevallier, T. (2013). Synthesis analysis of the temperature sensitivity of soil respiration from laboratory studies in relation to incubation methods and soil conditions. *Soil Biology and Biochemistry*, *58*, 115–126. https://doi.org/10.1016/j.soilbio.2012.11.012

Hirano, T., Kim, H., & Tanaka, Y. (2003). Long-term half-hourly measurement of soil CO2 concentration and soil respiration in a temperate deciduous forest. *Journal of Geophysical Research: Atmospheres*, *108*(D20). https://doi.org/10.1029/2003JD003766

Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017). The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls. *Annual Review of Ecology, Evolution and Systematics*, *48*(Volume 48, 2017), 419–445. https://doi.org/10.1146/annurev-ecolsys-112414-054234

Jian, J., Bailey, V., Dorheim, K., Konings, A. G., Hao, D., Shiklomanov, A. N., Snyder, A., Steele, M., Teramoto, M., Vargas, R., & Bond-Lamberty, B. (2022). Historically inconsistent productivity and respiration fluxes in the global terrestrial carbon cycle. *Nature Communications*, *13*(1), 1733. https://doi.org/10.1038/s41467-022-29391-5

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N., Manzon, J., Marchesi, R., Paredes, D., & Bond-Lamberty, B. (2021). A restructured and updated global soil respiration database (SRDB-V5). *Earth System Science Data*, *13*(2), 255–267. https://doi.org/10.5194/essd-13-255-2021

Jiang, J., Feng, L., Hu, J., Liu, H., Zhu, C., Chen, B., & Chen, T. (2024). Global soil respiration predictions with associated uncertainties from different spatio-temporal data

subsets. *Ecological Informatics*, *82*, 102777. https://doi.org/10.1016/j.ecoinf.2024.102777

Jobbágy, E. G., & Jackson, R. B. (2000). The Vertical Distribution of Soil Organic Carbon and its Relation to Climate and Vegetation. *Ecological Applications*, *10*(2), 423–436. https://doi.org/10.1890/1051-0761(2000)010%5B0423:TVDOSO%5D2.0.CO;2

Liu, K., Li, X., Wang, S., & Zhang, H. (2023). A robust gap-filling approach for European Space Agency Climate Change Initiative (ESA CCI) soil moisture integrating satellite observations, model-driven knowledge, and spatiotemporal machine learning. *Hydrology and Earth System Sciences*, *27*(2), 577–598. https://doi.org/10.5194/hess-27-577-2023

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, *21*(5), 1429–1442. https://doi.org/10.1890/09-1275.1

Maier, M., & Schack-Kirchner, H. (2014). Using the gradient method to determine soil gas flux: A review. *Agricultural and Forest Meteorology*, *192–193*, 78–95. https://doi.org/10.1016/j.agrformet.2014.03.006

Mariethoz, G., Linde, N., Jougnot, D., & Rezaee, H. (2015). Feature-preserving interpolation and filtering of environmental time series. *Environmental Modelling & Software*, *72*, 71–76. https://doi.org/10.1016/j.envsoft.2015.07.001

Millington, R. J., & Shearer, R. C. (1971). Diffusion in aggregated porous media. *Soil Science*, *111*(6), 372–378.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., & Stauch, V. J. (2007). Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, *147*(3), 209–232. https://doi.org/10.1016/j.agrformet.2007.08.011

Moldrup, P., Olesen, T., Yamaguchi, T., Schjønning, P., & Rolston, D. E. (1999). Modeling diffusion and reaction in soils: 9. The Buckingham-Burdine-Campbell equation for gas diffusivity in undisturbed soil. *Soil Science*, *164*(2), 75.

National Ecological Observatory Network (NEON). (2024a). *Barometric pressure (DP1.00004.001)*. National Ecological Observatory Network (NEON). https://doi.org/10.48443/RT4V-KZ04

National Ecological Observatory Network (NEON). (2024b). *Soil CO2 concentration (DP1.00095.001)*. National Ecological Observatory Network (NEON). https://doi.org/10.48443/E7GR-6G94

National Ecological Observatory Network (NEON). (2024c). *Soil physical and chemical properties, Megapit (DP1.00096.001)*. National Ecological Observatory Network (NEON). https://doi.org/10.48443/S6ND-Q840

National Ecological Observatory Network (NEON). (2024d). *Soil temperature (DP1.00041.001)*. National Ecological Observatory Network (NEON). https://doi.org/10.48443/Q24X-PW21

National Ecological Observatory Network (NEON). (2024e). *Soil water content and water salinity (DP1.00094.001)*. National Ecological Observatory Network (NEON). https://doi.org/10.48443/A8VY-Y813

Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown, K., & Vargas, R. (2017). The value of soil respiration measurements for interpreting and

modeling terrestrial carbon cycling. *Plant and Soil*, *413*(1), 1–25. https://doi.org/10.1007/s11104-016-3084-x

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). *Using Bayesian Model Averaging to Calibrate Forecast Ensembles*. https://doi.org/10.1175/MWR2906.1

Shao, J., Zhou, X., Luo, Y., Li, B., Aurela, M., Billesbach, D., Blanken, P. D., Bracho, R., Chen, J., Fischer, M., Fu, Y., Gu, L., Han, S., He, Y., Kolb, T., Li, Y., Nagy, Z., Niu, S., Oechel, W. C., … Zhang, J. (2015). Biotic and climatic controls on interannual variability in carbon fluxes across terrestrial ecosystems. *Agricultural and Forest Meteorology*, *205*, 11–22. https://doi.org/10.1016/j.agrformet.2015.02.007

Shao, P., Zeng, X., Moore, D. J. P., & Zeng, X. (2013). Soil microbial respiration from observations and Earth System Models. *Environmental Research Letters*, *8*(3), 034034. https://doi.org/10.1088/1748-9326/8/3/034034

Sihi, D., Gerber, S., Inglett, P. W., & Inglett, K. S. (2016). Comparing models of microbial–substrate interactions and their response to warming. *Biogeosciences*, *13*(6), 1733–1752. https://doi.org/10.5194/bg-13-1733-2016

Tang, J., Baldocchi, D. D., Qi, Y., & Xu, L. (2003). Assessing soil CO2 efflux using continuous measurements of CO2 profiles in soils with small solid-state sensors. *Agricultural and Forest Meteorology*, *118*(3), 207–220. https://doi.org/10.1016/S0168-1923(03)00112-6

Tang, J., Misson, L., Gershenson, A., Cheng, W., & Goldstein, A. H. (2005). Continuous measurements of soil respiration with and without roots in a ponderosa pine plantation in the Sierra Nevada Mountains. *Agricultural and Forest Meteorology*, *132*(3), 212–227. https://doi.org/10.1016/j.agrformet.2005.07.011

Zhang, R., Kim, S., Kim, H., Fang, B., Sharma, A., & Lakshmi, V. (2023). Temporal Gap-Filling of 12-Hourly SMAP Soil Moisture Over the CONUS Using Water Balance Budgeting. *Water Resources Research*, *59*(12), e2023WR034457. https://doi.org/10.1029/2023WR034457