

¹ **neonSoilFlux: An R Package for Continuous
2 Sensor-Based Estimation of Soil CO₂ Fluxes**

³ John Zobitz¹ Edward Ayres² Zoey Werbin³ Ridwan Abdi¹
⁴ Natalie Ashburner-Wright⁴ Lillian Brown⁴
⁵ Ryan Frink-Sobierajski⁴ Lajntxiag Lee¹ Dijonë Mehmeti¹
⁶ Christina Tran⁴ Ly Xiong¹ Naupaka Zimmerman^{4,5}

⁷ ¹ Augsburg University, 2211 Riverside Avenue, Minneapolis, MN 55454

⁸ ² National Ecological Observatory Network, Battelle, 1685 38th Street, Suite 100, Boulder,
⁹ CO 80301

¹⁰ ³ Boston University, 5 Cummington Street, Boston, MA 02215

¹¹ ⁴ University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117

¹² ⁵ University of Kansas, 1450 Jayhawk Boulevard, Lawrence, KS 66045

¹³ **Acknowledgments**

¹⁴ JZ acknowledges Kathleen O'Rourke for code development. NZ thanks technical staff at
¹⁵ USF for support with field gear assembly and shipping. We thank the NEON field staff
¹⁶ and assignable assets teams for facilitating each of the six NEON site visits. We are grateful

17 to LI-COR technical staff for helpful discussions about optimal soil chamber sampling methods.
18 This work was supported by NSF DEB grant #2017829 awarded to JZ, and NSF DEB grant
19 #2017860 awarded to NZ. This material is based in part upon work supported by the National
20 Ecological Observatory Network (NEON), a program sponsored by the U.S. National Science
21 Foundation (NSF) and operated under cooperative agreement by Battelle. We also thank the
22 reviewers and subject editor for their constructive feedback.

23 **Conflict of Interest Statements**

24 None of the authors have a financial, personal, or professional conflict of interest related to
25 this work.

26 **Author Contributions**

27 Conceptualization: JZ, NZ; Methodology: EA, JZ, NZ; Software: JZ, NZ, ZW, E A, DM, RA,
28 LX, LL; Validation: JZ, NZ; Formal Analysis: JZ, NZ, DM, RA, LX, LL; Investigation: JZ,
29 NZ, RF-S, CT, NA-W, LB; Resources: JZ, NZ; Data curation: JZ, NZ, DM, LX; Writing
30 – original draft: JZ, NZ; Writing – review and editing: JZ, NZ, ZW, EA, CT, DM, LX,;
31 Visualization: JZ, NZ, DM, RA, LX; Supervision: JZ; NZ; Project Administration: JZ; NZ;
32 Funding Acquisition: JZ; NZ

33 **Data Availability**

34 Anonymous field-collected data, `neonSoilFlux` calculated outputs, and manuscript-generating
35 code for peer review are provided as supplemental files. All will be made publicly available on

³⁶ Zenodo with a DOI upon publication.

³⁷ **1 Abstract**

³⁸ Accurate quantification of soil carbon fluxes is essential to reduce uncertainty in estimates of
³⁹ the terrestrial carbon sink. However, these fluxes vary over time and across ecosystem types
⁴⁰ and so it can be difficult to estimate them accurately across large scales. The flux gradient
⁴¹ method estimates soil carbon fluxes using co-located measurements of soil CO₂ concentration,
⁴² soil temperature, soil moisture, and other soil properties. The National Ecological Observatory
⁴³ Network (NEON) provides such data across 20 ecoclimatic domains spanning the continental
⁴⁴ U.S., Puerto Rico, Alaska, and Hawai‘i. We present an R software package (`neonSoilFlux`)
⁴⁵ that acquires soil environmental data to compute half-hourly soil carbon fluxes for each soil
⁴⁶ replicate plot at a given terrestrial NEON site. To assess the computed fluxes, we visited six
⁴⁷ focal NEON sites and measured soil carbon fluxes using a closed-dynamic chamber approach.
⁴⁸ Outputs from the `neonSoilFlux` showed agreement with measured fluxes (R^2 between mea-
⁴⁹ sured and `neonSoilFlux` outputs ranging from 0.04 to 0.81 depending on calculation method
⁵⁰ used); measured outputs generally fell within the range of calculated uncertainties from the
⁵¹ gradient method. Calculated fluxes from `neonSoilFlux` aggregated to the daily scale exhibited
⁵² expected site-specific seasonal patterns. While the flux gradient method is broadly effective,
⁵³ its accuracy is highly sensitive to site-specific inputs, including the extent to which gap-filling
⁵⁴ techniques are used to interpolate missing sensor data and to estimates of soil diffusivity and
⁵⁵ moisture content. Future refinement and validation of `neonSoilFlux` outputs can contribute
⁵⁶ to existing databases of soil carbon flux measurements, providing near real-time estimates of
⁵⁷ a critical component of the terrestrial carbon cycle.

58 **1.1 Keywords**

59 Soil carbon, carbon dioxide, flux gradient, carbon cycle, field validation, soil respiration, ecosys-
60 tem variability, diffusion

61 **2 Data for peer review**

62 Anonymous field-collected data, `neonSoilFlux` calculated outputs, and manuscript-generating
63 code for peer review are provided as supplemental files. All will be made publicly available on
64 Zenodo with a DOI upon publication.

65 **3 Introduction**

66 Soils contain the planet's largest reservoir of terrestrial carbon (Jobbágy & Jackson, 2000). A
67 critical component of this reservoir is soil organic matter, the accumulation of which is influ-
68 enced by biotic factors such as above-ground plant inputs (Jackson et al., 2017). These inputs
69 in turn are influenced by environmental factors such as growing season length, temperature,
70 and moisture (Desai et al., 2022), which also affect the breakdown of soil organic matter and its
71 return to the atmosphere. Across heterogeneous terrestrial landscapes, the interplay between
72 these biotic and abiotic factors influence the size of the soil contribution to the terrestrial
73 carbon sink (Friedlingstein et al., 2025). However, the heterogeneity of these processes across
74 diverse ecosystems in the context of rapid environmental change leads to large uncertainty
75 about the magnitude of this sink in the future, and thus there remains a pressing need to
76 quantify changes in soil carbon pools and fluxes across scales.

77 Ecological observation networks such as the United States' National Ecological Observatory
78 Network (NEON) and others (e.g. the globally-distributed FLUXNET or the European Inte-
79 grated Carbon Observation System) present a significant advancement in the nearly continuous
80 observation of biogeochemical processes at the continental scale. Notably, at 47 terrestrial sites
81 across the continental United States that span 20 ecoclimatic domains, NEON provides half-
82 hourly measurements of soil CO₂ concentration, temperature, and moisture at different vertical
83 depths. Each of these NEON sites also encompasses measurements of the cumulative sum of all
84 ecosystem carbon fluxes in an airshed using the eddy covariance technique (Baldocchi, 2014).
85 Soil observations provided by NEON are on the same timescale and standardized with eddy co-
86 variance measurements from FLUXNET. These types of nearly continuous observational data
87 (NEON and FLUXNET) can be used to reconcile differences between model-derived or data-
88 estimated components of ecosystem carbon flux (Jian et al., 2022; Luo et al., 2011; Phillips et
89 al., 2017; J. Shao et al., 2015; P. Shao et al., 2013; Sihl et al., 2016).

90 Estimated or observed soil carbon fluxes are a key metric for understanding change in soil
91 carbon pools over time (Bond-Lamberty et al., 2024). A soil carbon flux to the atmosphere
92 (F_S , units $\mu\text{mol m}^{-2} \text{ s}^{-1}$), represents the aggregate process of transfer of soil CO₂ to the
93 atmosphere from physical and biological processes (e.g. diffusion and respiration). Soil carbon
94 fluxes can be assumed to encompass soil carbon respiration from autotrophic or heterotrophic
95 sources (Davidson et al., 2006) and modeled with a exponential Q_{10} paradigm (Bond-Lamberty
96 et al., 2004; Chen & Tian, 2005; Hamdi et al., 2013).

97 One common method by which F_S is measured in the field is through the use of soil chambers
98 in a closed, well-mixed system (Norman et al., 1997) with headspace trace gas concentrations
99 measured with an infrared gas analyzer (IRGA). F_S can also be estimated from soil CO₂
100 measurements at different depths in the soil using the flux-gradient method (Maier & Schack-
101 Kirchner, 2014). Closed-chamber IRGA measurements, while being the most common method,

102 require either frequent in-person site visits or expensive and fragile automated systems. The
103 potential of the gradient method is that fluxes can be estimated from continuous data recorded
104 by robust solid-state sensors. The flux-gradient method is an approach that uses conservation
105 of mass to calculate flux at a vertical soil depth z at steady state by applying Fick's law of
106 diffusion. A simplifying assumption for the flux-gradient method is that there is no mass trans-
107 fer in the other spatial dimensions x and y (Maier & Schack-Kirchner, 2014). The diffusivity
108 profile, a key component of this calculation, varies across the soil depth as a function of soil
109 temperature, soil volumetric water content, atmospheric air pressure, and soil bulk density
110 (Millington & Shearer, 1971; Moldrup et al., 1999; Sallam et al., 1984).

111 Databases such as the Soil Respiration Database (SRDB) or the Continuous Soil Respiration
112 Database (COSORE) add to the growing network of resources for making collected observa-
113 tions of soil fluxes available to other researchers (Bond-Lamberty, 2018; Bond-Lamberty et
114 al., 2020; Bond-Lamberty & Thomson, 2010; Jian et al., 2021; Jiang et al., 2024). However,
115 these databases currently encompass primarily direct soil measurements of fluxes (i.e. those
116 using methods like the closed-chamber method described above). Currently, NEON provides
117 all measurements to calculate F_S from Fick's law, but soil flux as a derived data product was
118 descoped from the initial network launch due to budget constraints (Berenbaum et al., 2015).
119 Deriving estimates of F_S using continuous sensor data across NEON sites thus remains a high
120 priority.

121 This study describes an R software package, `neonSoilFlux`, that computes a standardized
122 estimate of F_S at all terrestrial NEON sites using the flux-gradient method. Using direct
123 chamber-based field observations of soil carbon dioxide flux from a subset of terrestrial NEON
124 sites spanning six states, we provide a direct validation of F_S from `neonSoilFlux`.

125 Key objectives of this study are to:

- 126 1. Apply the flux-gradient method to estimate soil CO₂ flux from continuous sensor mea-
127 surements across six NEON sites.
- 128 2. Benchmark estimated soil carbon fluxes against field measurements (e.g., direct chamber
129 measurements of soil flux).
- 130 3. Identify sources of error in the flux-gradient approach across diverse sites in order to
131 guide future work.

132 4 Materials and Methods

133 4.1 Field methods

134 4.1.1 Focal NEON Sites

135 In order to acquire field data to validate model predictions of flux, we selected six terrestrial
136 NEON sites for analysis. We conducted roughly week-long field measurement campaigns at
137 these sites, which span a range of environmental gradients and terrestrial domains (Table 1).
138 SJER, SRER, and WREF were visited during May and June of 2022, and WOOD, KONZ,
139 and UNDE during May and June of 2024.

140 4.1.2 Soil collar placement

141 Either one (2022 sampling campaign) or two (2024 sampling campaign) PVC soil collars (20.1
142 cm inside diameter) were installed in close proximity to the permanent NEON soil sensors at
143 each site (Figure 1). As instruments in the NEON soil sensor arrays can occasionally break
144 down or stop working, the specific soil plot where we made measurements was chosen at each
145 site in consultation with NEON staff to maximize likelihood of quality soil sensor measurements

¹⁴⁶ during the duration of the IRGA measurements. The plot selected at each site (out of the 5 in
¹⁴⁷ each replicate array at each site) are presented in the last column of Table 1. After installation,
¹⁴⁸ collar(s) were left to equilibrate for approximately 24 hours prior to any measurements being
¹⁴⁹ taken.

¹⁵⁰ **4.1.3 Infrared gas analyzer measurements of soil CO₂ flux**

¹⁵¹ In 2022, we then made measurements of flux on an hourly interval for 8 hours each day.
¹⁵² Measurements were taken from roughly 8 am to 4 pm, with the time interval selected to
¹⁵³ capture the majority of the diurnal gradient of soil temperature each day. These measurements
¹⁵⁴ were made using a LI-6800 infrared gas analyzer instrument (LI-COR Environmental, Lincoln,
¹⁵⁵ NE) fitted with a soil chamber attachment (attachment 6800-09). In 2024, we again used the
¹⁵⁶ same LI-6800 instrument, but made half-hourly measurements over an approximately 8 hour
¹⁵⁷ period. In addition, in 2024 we also installed a second collar and used a second instrument, an
¹⁵⁸ LI-870 CO₂ IRGA, connected to an automated robotic chamber (LI-COR chamber 8200-104)
¹⁵⁹ controlled by an LI-8250 multiplexer to make automated measurements. The multiplexer was
¹⁶⁰ configured to take half-hourly measurements 24 hours a day for the duration of our sampling
¹⁶¹ bout at each site. Each instrument was paired with a soil temperature and moisture probe
¹⁶² (Stevens HydraProbe, Stevens Water, Portland, OR) that was used to make soil temperature
¹⁶³ and moisture measurements concurrent with the CO₂ flux measurements. Chamber volumes
¹⁶⁴ were set by measuring collar offsets at each site. System checks were conducted daily for the
¹⁶⁵ LI-6800 and weekly for the LI-8250. Instruments were factory calibrated before each field
¹⁶⁶ season.

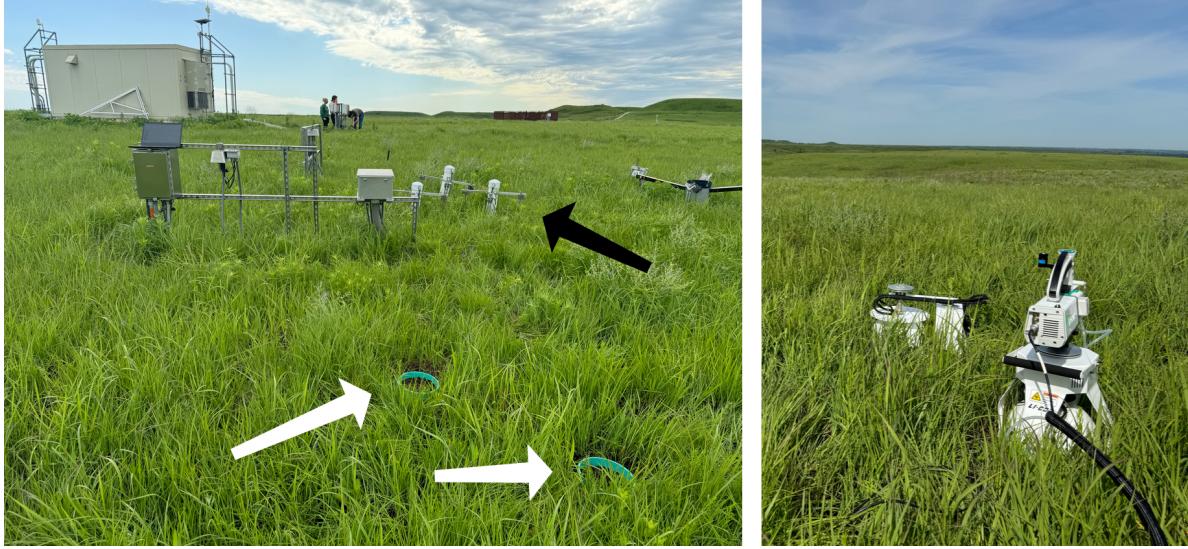


Figure 1: Spatial layout of field sampling using a closed-dynamic chamber setup at a representative NEON site (KONZ). Left image shows collars (white arrows) and permanent soil sensor installation (black arrow) and right image shows the LI-6800 (foreground) and LI-8200-104 (background) instruments placed on the collars.

Table 1: Listing of NEON sites studied for field work and analysis. Site refers to NEON site codes: Santa Rita Experimental Range (SRER), San Joaquin Experimental Range (SJER), Wind River Experimental Forest (WREF), Chase Lake National Wildlife Refuge (WOOD), Konza Prairie Biological Station (KONZ), and the University of Notre Dame Environmental Research Center (UNDE). Location is reported in decimal degrees of latitude and longitude. Other abbreviations include Mean Annual Temperature (MAT); \bar{T}_S : average soil temperature during field measurements; \bar{SWC} : average soil water content during field measurements. Dates refer to field measurement dates for each site. Plot refers to the particular location in the soil sensor array (denoted as HOR by NEON) where field measurements were made.

Site	Location	Ecosystem	MAT	\bar{T}_S	MAP	\bar{SWC}	Dates	Plot
SRER	31.91068, -110.83549	Shrubland	19.3 °C	47.6 °C	346 mm	4.0%	May 29– June 1 2022	004
SJER	37.10878, -119.73228	Oak woodland	16.4 °C	41.7 °C	540 mm	1.2%	June 1–4 2022	005
WREF	45.82049, -121.95191	Evergreen forest	9.2 °C	15.3 °C	2225 mm	27.2%	June 7–9 2022	001
WOOD	47.1282, -99.241334	Restored prairie	4.9 °C	14.9 °C	495 mm	14.9%	June 3–9 2024	001
KONZ	39.100774, -96.563075	Tallgrass prairie	12.4 °C	23.4 °C	870 mm	23.4%	May 29– June 1 2024	001

Table 1: Listing of NEON sites studied for field work and analysis. Site refers to NEON site codes: Santa Rita Experimental Range (SRER), San Joaquin Experimental Range (SJER), Wind River Experimental Forest (WREF), Chase Lake National Wildlife Refuge (WOOD), Konza Prairie Biological Station (KONZ), and the University of Notre Dame Environmental Research Center (UNDE). Location is reported in decimal degrees of latitude and longitude. Other abbreviations include Mean Annual Temperature (MAT); \bar{T}_S : average soil temperature during field measurements; \bar{SWC} : average soil water content during field measurements. Dates refer to field measurement dates for each site. Plot refers to the particular location in the soil sensor array (denoted as HOR by NEON) where field measurements were made.

Site	Location	Ecosystem	MAT	\bar{T}_S	MAP	\bar{SWC}	Dates	Plot
UNDE	46.23391, -89.537254	Deciduous forest	4.3 °C	13.0 °C	802 mm	13.0%	May 22–25 2024	004

167 4.1.4 Post-collection processing of field data

168 We used LI-COR SoilFluxPro software (v 5.3.1) to assess the data after collection and to inform
 169 sampling parameters. We checked appropriateness of dead band and measurement durations
 170 using built-in evaluation tools. Based on this, the deadband period was set for 30-40 seconds,
 171 depending on the site, and the measurement duration was 180 seconds with a 30 second pre-
 172 purge and a 30 second post-purge at most sites, and a 90 second pre- and post-purge at sites
 173 with higher humidity due to recent precipitation events. We also assessed the R^2 of linear and
 174 exponential model fits to measured CO₂ to verify measurement quality.

175 4.2 neonSoilFlux R package

176 We developed an R package called `neonSoilFlux` (Zobitz et al., 2024) to compute half-hourly
 177 soil carbon fluxes and uncertainties from NEON data. The objective of the `neonSoilFlux`
 178 package is a unified workflow (Figure 2) for soil data acquisition and analysis that supplements
 179 the existing `neonUtilities` data acquisition R package (Lunch et al., 2025).

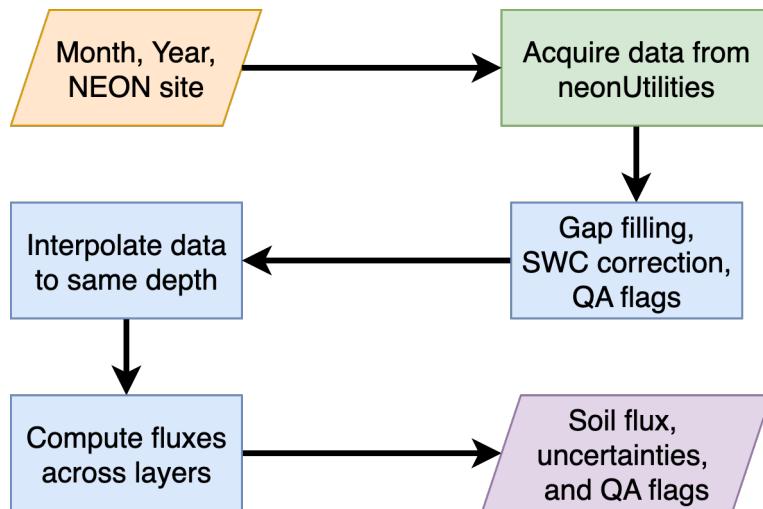


Figure 2: Diagram of `neonSoilFlux` R package. For a given month, year, and NEON site (orange parallelogram), the package acquires all relevant data to compute F_S using the `neonUtilities` R package (green rectangle). Data are gap-filled according to reported QA flags and adjusted for changes in soil water content (SWC) calibration coefficients, then interpolated to the same measurement depth before computing the soil flux, uncertainties, and final QA flags (blue rectangles). The package reports the associated soil flux, uncertainties, and quality assurance (QA) flags for the user (purple parallelogram).

180 At a given NEON site there are five replicate soil plots, each with measurements of soil
 181 CO_2 concentration, soil temperature, and soil moisture at different depths (Figure 3). The
 182 `neonSoilFlux` package acquires measured soil CO_2 concentration (National Ecological Ob-
 183 servatory Network (NEON), 2024b), soil temperature (National Ecological Observatory Net-
 184 work (NEON), 2024d), soil water content (National Ecological Observatory Network (NEON),
 185 2024e), barometric pressure from the nearby tower (National Ecological Observatory Network
 186 (NEON), 2024a), and soil properties (e.g. bulk density) (National Ecological Observatory Net-
 187 work (NEON), 2024c) from a range of different NEON data products. The static soil properties
 188 were collected by NEON staff from a nearby soil pit during initial site characterization and
 189 are assumed to be constant at each site. A soil flux calculation is computed at each replicate
 190 soil plot.

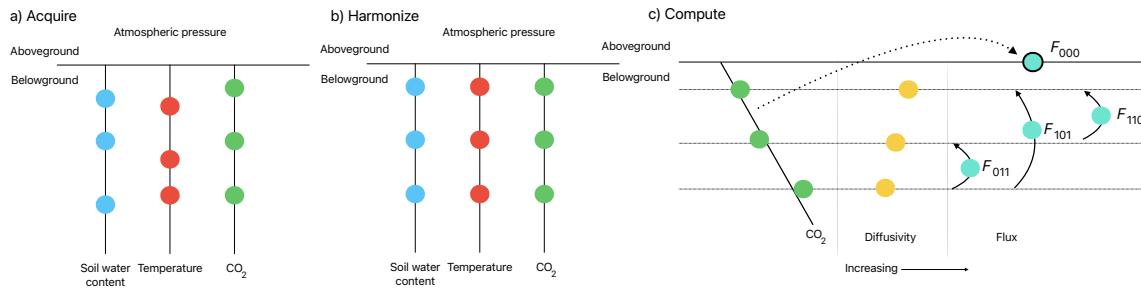


Figure 3: Model diagram of the data workflow for the `neonSoilFlux` R package. a) Acquire: Data are obtained for a given NEON location and horizontal sensor location, which includes soil water content, soil temperature, CO_2 concentration, and atmospheric pressure. All data are screened for quality assurance; if gap-filling of missing data occurs, it is flagged for the user. b) Harmonize: Any belowground data are then harmonized to the same depth as CO_2 concentrations using linear regression. c) Compute: 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 layers the flux is computed across ($i =$ closest to surface, $k =$ deepest). F_{000} represents a flux estimate where the gradient dC/dz is the slope of a linear regression of CO_2 with depth.

191 The workflow to compute a value of F_S with `neonSoilFlux` consists of three primary steps,
 192 illustrated in Figure 3. First, NEON data are acquired for a given site and month via the

193 `neonUtilities` R package (yellow parallelogram and green rectangle in Figure 2 and Panel
194 a in Figure 3). Acquired environmental data can be exported to a comma separated value
195 file for additional analysis. Quality assurance (QA) flags are reported as an indicator variable.
196 Since the calibration coefficients on the soil water content sensors have changed over time
197 (National Ecological Observatory Network (NEON), 2024e), raw sensor measurements were
198 back-calculated and soil-specific calibrations were applied following Ayres et al. (2024) to
199 generate a consistent time series at each measurement location.

200 The second step is harmonizing the data to compute soil fluxes across soil layers. This step
201 consists of three different actions (blue rectangles in Figure 2 and Panel b in Figure 3). If a
202 given observation by NEON is reported as not passing a quality assurance check, we applied
203 a gap filling method to replace that measurement with its monthly mean at that same depth
204 (Section 4.2.1). Belowground measurements of soil water and soil temperature are then inter-
205 polated to the same depth as soil CO₂ measurements. The diffusivity (Section 4.2.2) and soil
206 flux across different soil layers (Section 4.2.3) are then computed.

207 The third and final step is computing a surface soil flux through extrapolation to the sur-
208 face (purple parallelogram in Figure 2 and Panel c in Figure 3). Uncertainty on a soil flux
209 measurement is computed through quadrature. An aggregate quality assurance (QA) flag
210 for each environmental measurement is also reported, representing if any gap-filled measure-
211 ments were used in the computation of a soil flux. Within the soil flux-gradient method,
212 several different approaches can be used to derive a surface flux (Maier & Schack-Kirchner,
213 2014); the `neonSoilFlux` package reports four different possible values for soil surface flux
214 (Section 4.2.3) for each of two different methods of diffusivity estimation, for a total of eight
215 estimates of flux.

216 **4.2.1 Gap-filling routine**

217 NEON reports QA flags as binary values for each measurement and half-hourly interval. For
218 a given half-hour, if any input variable (soil CO₂ concentration, soil temperature, or soil
219 moisture) at depth z is flagged, computation of F_S is not possible. To address this, flagged
220 measurements and their uncertainties were replaced with a bootstrapped monthly mean (\bar{m})
221 and monthly standard deviation (\bar{s}) (Efron & Tibshirani, 1994).

222 For each month, depth z , and variable, we computed bootstrapped estimates of \bar{m} and \bar{s}
223 from the vectors of unflagged measurements (**m**), reported standard errors (σ), and the 95%
224 confidence interval (ϵ , or expanded uncertainty; Farrance & Frenkel (2012)). We also defined
225 a bias vector $\mathbf{b} = \sqrt{\epsilon^2 - \sigma^2}$, which quantifies the spread of uncertainty in a given period and
226 is incorporated into \bar{m} .

227 From these, 5000 bootstrap samples were generated for **m**, σ , and **b**. For each sample
228 (m_k, b_k, σ_k), we generated a vector **n** (length $N = 5000$) by drawing from a normal dis-
229 tribution with mean $m_k + b_k$ and standard deviation σ_k . The sample mean and standard
230 deviation were then computed from **n**. The resulting distributions of sample means and
231 sample standard deviations provided the bootstrapped monthly mean (\bar{m}) and standard error
232 (\bar{s}) respectively.

233 This gap-filling procedure provides a consistent treatment across all data streams. However,
234 alternative approaches may be better suited for longer gaps (e.g., correlations with other
235 NEON measurement levels or soil plots) or for variable-specific conditions. We discuss the
236 effect of gap-filling on our results in Section 6.1.

²³⁷ **4.2.2 Soil diffusivity**

²³⁸ Soil diffusivity D_a at a given measurement depth is the product of the diffusivity in free air
²³⁹ $D_{a,0}$ ($\text{m}^2 \text{ s}^{-1}$) and the tortuosity ξ (no units) (Millington & Shearer, 1971).

²⁴⁰ We compute $D_{a,0}$ with Equation 1:

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

²⁴¹ where T_i is soil temperature ($^\circ\text{C}$) at depth i (National Ecological Observatory Network
²⁴² (NEON), 2024d) and P surface barometric pressure (kPa) (National Ecological Observatory
²⁴³ Network (NEON), 2024a).

²⁴⁴ Previous studies by Sallam et al. (1984) and Tang et al. (2003) demonstrated the sensitivity
²⁴⁵ of modeled F_S depending on the tortuosity model (ξ) used to compute diffusivity. At low
²⁴⁶ soil water content, the choice of tortuosity model can lead to order-of-magnitude differences
²⁴⁷ in D_a , which in turn affect modeled F_S . The `neonSoilFlux` package currently includes two
²⁴⁸ approaches to calculate ξ , representing the range of tortuosity behavior reported in Sallam et
²⁴⁹ al. (1984).

²⁵⁰ The first approach is the Millington-Quirk model (Millington & Shearer, 1971), in which
²⁵¹ tortuosity depends on both porosity and soil water content:

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

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

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

255 In Equation 3, ρ_m is the particle density of mineral soil (2.65 g cm^{-3}), ρ_s the soil bulk density
256 (g cm^{-3}) excluding coarse fragments greater than 2 mm (National Ecological Observatory
257 Network (NEON), 2024c), and f_V is a site-specific value that accounts for the proportion of
258 soil fragments between 2-20 mm. Soil fragments greater than 20 mm were not estimated due
259 to limitations in the amount of soil that can be analyzed (National Ecological Observatory
260 Network (NEON), 2024c). We assume that rock fragments contain no internal pores.

261 The Millington-Quirk model assumes ξ is modulated by the amount of fluid saturation in
262 soil pores (Millington & Shearer, 1971). In contrast, the Marshall model (Marshall, 1959)
263 expresses tortuosity as only a function of porosity ($\xi = \phi^{1.5}$), with ϕ defined from Equation
264 3. The Marshall model is independent of soil water content and assumes tortuosity is only
265 governed by soil structure. The `neonSoilFlux` package allows users to choose the tortuosity
266 model most appropriate for site-specific conditions and research goals.

267 **4.2.3 Soil flux computation**

268 We applied Fick's law (Equation 4) to compute the soil flux F_{ij} ($\mu\text{mol m}^{-2} \text{ s}^{-1}$) across two
269 soil depths i and j :

$$F_{ij} = -D_a \frac{dC}{dz} \quad (4)$$

where D_a is the diffusivity ($\text{m}^2 \text{ s}^{-1}$) and $\frac{dC}{dz}$ is the gradient of CO_2 molar concentration ($\mu\text{mol m}^{-3}$, so the gradient has units of $\mu\text{mol m}^{-3} \text{ m}^{-1}$). 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 4 across different soil depths. Each site had three measurement layers, so we denote the flux as a three-digit subscript F_{ijk} with indicator variables i , j , and k indicate if a given layer was used (written in order of increasing depth), according to the following:

- F_{000} is a surface flux estimate using the intercept of the linear regression of D_a with depth and the slope from the linear regression of CO_2 with depth (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_{110} , F_{011} 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_{101} is a surface flux estimate using fluxes estimated at the shallowest and deepest measurement layers (F_{110} , F_{011}). These computed fluxes then form the basis of additional linear extrapolation to the surface. For example, 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.

Uncertainty in all F_{ijk} is computed through quadrature (Taylor, 2022).

292 **4.3 Post processing evaluation**

293 Following collection of field measurements and calculation of the soil fluxes from `neonSoilFlux`
294 package, we compared measured F_S based on closed-dynamic chamber measurements with the
295 LI-COR instruments to a given soil flux calculation from `neonSoilFlux` for each site and flux
296 computation method and quantified the relationship statistically (R^2). Finally, for a half-
297 hourly interval we also computed a *post hoc* diffusivity (D_a) using the LI-COR flux along
298 with the CO_2 surface gradient reported by NEON using the measurement levels closest to the
299 surface.

300 **5 Results**

301 **5.1 Concordance between modelled and measured soil CO_2 flux**

302 The sites we visited ranged substantially in both their annual average temperature and precip-
303 itation as well as their biome type (Table 2). These differences also influenced the wide range
304 of observed flux rates across sites.

305 The timeseries of the measured fluxes from the LI-COR 6800 and 870/8250 were compared
306 to modeled soil fluxes from the `neonSoilFlux` R package (Figure 4). We also assessed year-
307 long estimated flux time series and compared those to field measurements made at each site
308 (Figure 5). Results are reported in local time. Where applicable, sites are displayed from left
309 to right by increasing soil temperature (Table 1). Positive values of the flux indicate that there
310 is a flux moving towards the surface. Overall, with the exception of SRER (discussed later) the
311 computed fluxes determined using a variety of plausible methods spanned the field-measured
312 fluxes, but the specific flux-gradient method that best approximated field measurements varied
313 by site.

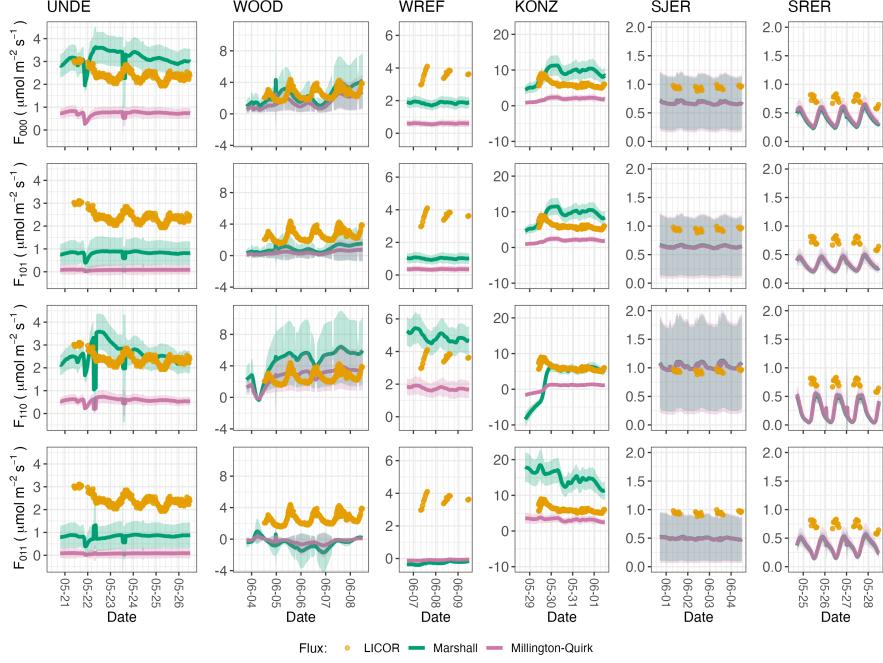


Figure 4: Timeseries of soil surface flux (F_S) from field-measured (yellow lines) and modeled soil fluxes (green or purple lines) by the `neonSoilFlux` R package. Fluxes from the `neonSoilFlux` R package are separated by the diffusivity model used (Millington-Quirk or Marshall, Section 4.2.2). Individual vertical axis labels in the first column represent the measurement levels where the flux-gradient approach is applied (Section 4.2.3). Ribbons for modeled soil fluxes represent ± 1 standard deviation. Results are reported in local time. WREF, SJER, and SRER were sampled in 2022, and UNDE, WOOD, and KONZ were sampled in 2024. Sites (columns) are arranged from left to right in terms of increasing mean annual temperature.

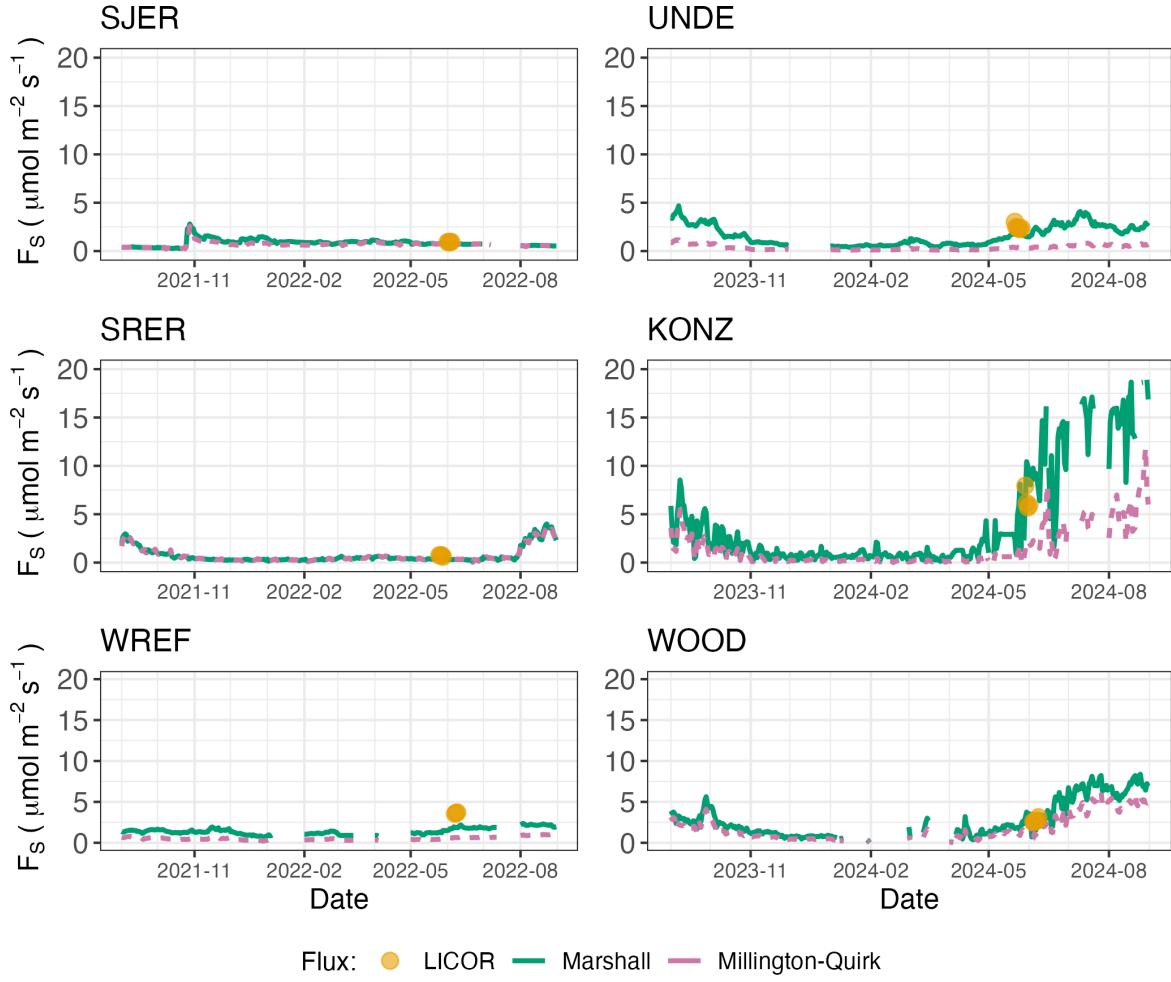


Figure 5: Timeseries of both daily-averaged field F_S (yellow circles) and daily ensemble averaged soil fluxes (average of F_{000} , F_{101} , F_{011} , F_{110} , Section 4.2.3) by the `neonSoilFlux` R package, separated by the diffusivity model used (green or purple lines, Millington-Quirk or Marshall, Section 4.2.2).

Table 2: Summary of measured soil characteristics and flux results from field measurements across six NEON sites using a LI-COR 6800 (LI-870/8250 measurements omitted to enable direct comparability) via the closed-dynamic chamber method. Numeric values for soil CO₂ flux, soil temperature, and volumetric soil water content (VSWC) are the mean and standard deviation of field measurements at each site.

Site	Flux μmol m ⁻² s ⁻¹	Soil temp °C	VSWC cm ³ cm ⁻³	n
UNDE	2.55 ± 0.26	14.33 ± 0.77	0.33 ± 0.02	61
WOOD	3.02 ± 0.4	16.01 ± 1.54	0.28 ± 0.01	53
WREF	3.62 ± 0.3	15.34 ± 1.76	0.27 ± 0.06	21
KONZ	6.35 ± 0.97	27.28 ± 4.14	0.37 ± 0.01	44
SJER	0.94 ± 0.02	41.68 ± 11.22	0.01 ± 0.01	32
SRER	0.72 ± 0.09	47.64 ± 7.46	0.04 ± 0.01	32

³¹⁴ We calculated a statistical relationship between the various estimates of soil flux computed by
³¹⁵ `neonSoilFlux` and the field-measured fluxes within daily interval periods. Statistics for these
³¹⁶ comparisons are reported in Figure 6, which also shows how these fall relative to a 1:1 line.

³¹⁷ 5.2 Effects of method choice on diffusivity estimates

³¹⁸ In four of six field sites, the *post hoc* D_a estimate fell roughly between the two diffusion
³¹⁹ estimation methods; however this was less the case in the two driest sites, SJER and SRER
³²⁰ (Table 1), where the field estimate of diffusivity was either lower or higher than both of the
³²¹ other methods (Figure 7).

³²² 6 Discussion

³²³ This study presents a unified data science workflow to efficiently process automated measure-
³²⁴ ments of belowground soil CO₂ concentrations, soil water content, and soil temperature to
³²⁵ infer estimates of soil surface CO₂ effluxes through application of Fick's Law (Equation 4).

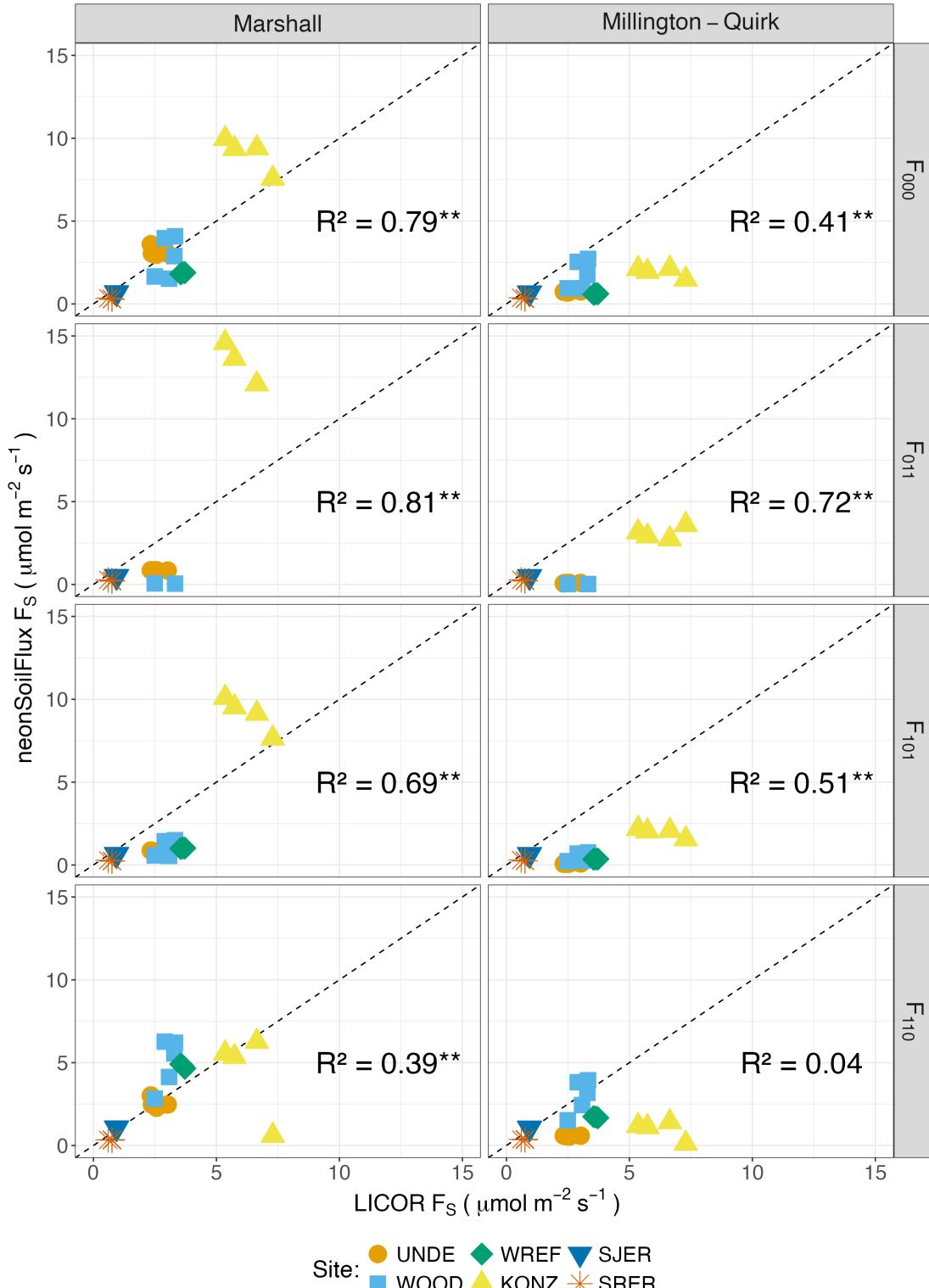


Figure 6: Statistical comparison between measured fluxes at each NEON site with fluxes reported by `neonSoilFlux` with the different flux calculation approaches and diffusivity calculations applied. Points are daily averages and LICOR F_S values are from the 6800 instrument only, for consistency. The dotted line represents a 1:1 relationship, and the reported R^2 quantifies the relationship between field-measured and `neonSoilFlux` estimated fluxes. * = significance at the 5% level, ** = significance at the 1% level. The low-value outlier from KONZ in the F_{110} Marshall plot is an example of the effect of inverted CO₂ gradients causing an estimated flux to be negative, bringing down the daily mean, which later resolved as the soils dried back out.

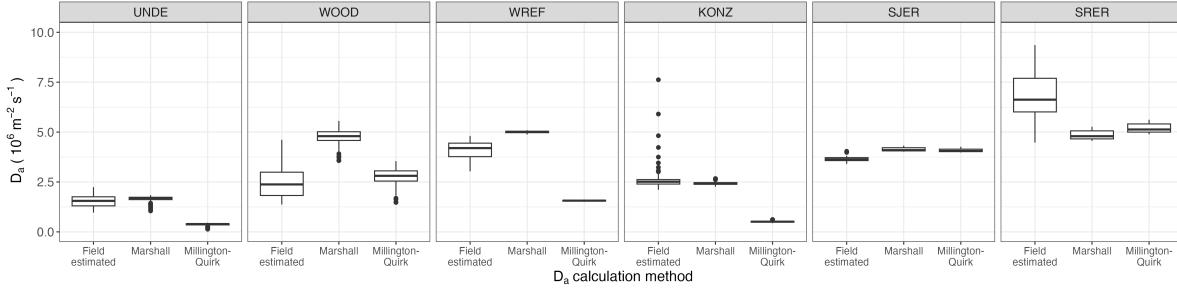


Figure 7: Distribution of diffusivity (D_a) at each study site. Values of D_a were provided by the `neonSoilFlux` package, computed from the Millington-Quirk or Marshall models (Section 4.2.2). A post-hoc estimate of diffusivity (labeled as “Field estimated”) was computed through the field measured flux (Figure 4), divided by the CO_2 gradient from the measurement levels closest to the soil surface, as reported by NEON. We only used F_S measured by the LICOR 6800 at all sites to standardize comparisons.

326 Our core goals in this study were: (1) to generate estimates of soil flux from continuous soil
 327 sensor data at terrestrial NEON sites using the flux-gradient method and then (2) to compare
 328 those estimates to field-measured fluxes based on the closed chamber approach at six NEON
 329 focal sites. We discuss our progress toward these core goals through (1) an overall evaluation
 330 of the flux-gradient approach (and uncertainty calculation) and (2) site-specific evaluation of
 331 differences in estimated vs measured fluxes.

332 **6.1 General evaluation of flux-gradient approach**

333 Key assumptions of the flux-gradient approach are that CO_2 concentrations increase through-
 334 out the soil profile such that the highest concentrations are observed in the deepest layers. Ad-
 335 ditionally, field flux measurements should correlate with F_{000} because they represent surface
 336 fluxes. Periods where this gradient condition are not met generally are connected to processes
 337 that occur during soil wetting events, where more shallow soil layers produce higher concentra-
 338 tions of CO_2 due to microbial respiration pulses following rewetting. This effect is likely to be
 339 largest at sites with rich organic soils (e.g. KONZ). Based on this reasoning, in these types of

340 situations we would *a priori* expect F_{011} (deepest layers) $\leq F_{101} \leq F_{110}$ (shallow layers) \leq
341 F_{000} (all layers) because the previous flux estimates rely primarily on CO₂ concentrations at
342 deeper depths, and could miss high concentrations of CO₂ produced in shallower layers.

343 When modeling soil respiration, typically a non-linear response function that also considers
344 soil type is used (Bouma & Bryla, 2000; Yan et al., 2016, 2018). For the `neonSoilFlux`
345 package, soil type is connected to the measurement of bulk density, which was characterized
346 at each NEON site. This bulk density estimate is based on replicate samples collected from
347 the site megapit at a subset of soil horizons, with an estimated uncertainty of $\pm 5\%$ (National
348 Ecological Observatory Network (NEON), 2024c). Coarse fragment estimates also have very
349 large uncertainties, but because the volume fraction tends to be low in surface soils it is unlikely
350 to contribute much additional flux uncertainty.

351 Our results suggest that the most important way to improve reliability of the flux estimate is
352 to reduce the usage of gap-filled data. The current approach to gap filling in `neonSoilFlux`
353 uses monthly mean data to gap fill—this approach decreases the ability of the estimate to be
354 responsive to short-term pulses that occur with rapid weather shifts. Four sites (KONZ, SRER,
355 WREF, and UNDE) had more than 75% of half-hourly periods with no-gap filled measurements
356 (Figure S1, Supplementary Information). Two sites (SJER and WOOD) had more than 75%
357 of half-hourly intervals with just one gap-filled measurement. The large uncertainty evident
358 in Figure 4 for estimates from WOOD and SJER are thus due in part to the gap-filling used
359 in these sites (Figure S1). While we did not need to use gap-filled measurements to compute
360 the flux at WREF, field data collection occurred following a severe rainstorm, with soils at the
361 beginning of the sampling week near their water holding capacity. In general, we recommend
362 that whenever possible, knowledge of local field conditions should influence analysis decisions
363 in addition to any QA filtering protocols in the `neonSoilFlux` package.

364 We recognize that this gap-filling approach may lead to gap-filled values that are quite different

from the actual values, such as an underestimate of soil moisture following rain events. Further extensions of the gap filling 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). Additionally, since the deepest temperature and soil moisture sensors are located below the deepest CO₂ sensors at NEON sites, it is possible that excluding these deeper layers from consideration prior to analysis would lead to a reduced need for gap filling. Future iterations of the `neonSoilFlux` package may incorporate this as an option. The current gap-filling routine provides a consistent approach that can be applied to each data stream, but further work may explore alternative gap-filling approaches.

6.2 Evaluation of flux-gradient approach at each site

Derived results from the `neonSoilFlux` package have patterns that are broadly consistent with those directly measured in the field (Figure 4 and Figure 5), even though statistical comparisons between the field-measured and `neonSoilFlux` values were quite variable (e.g. R^2 ranging from 0.04 to 0.81; Figure 6). One advantage of the `neonSoilFlux` package is its ability to calculate fluxes across different soil depths (Figure 3), which allows for additional site-specific customization. We believe the package can provide a useful baseline estimate of soil fluxes that can always be complemented through additional field measurements.

The six locations studied provide a range of case studies that suggest different considerations may apply to different sites when applying the flux-gradient method. For example, the Santa Rita Experimental Range (SRER) is a desert site characterized by sandy soil, which also was the location of the highest field soil temperatures that we observed (Table 2). At SRER the flux across the top two layers (F_{110}) produced a pattern of soil flux most consistent with the observed field data. The remaining methods F_{101} , F_{011} , or F_{000} are derived from information taken from the deepest layer, which seems to have been decoupled from the surface layers both

389 in terms of temperature and CO₂ concentration. This may be a general circumstance where
390 there are large diurnal temperature extremes that rapidly change during the course of a day
391 and overnight, leading to lags in the timing of when temperature increases propagate down to
392 deeper soil layers.

393 Immediately prior to our visit to Konza Prairie (KONZ), that site that experienced a significant
394 rain event that led to wet soils that gradually dried out over the course of our time there.
395 This pulse of precipitation increased the soil CO₂ concentration at the top layer above the
396 concentrations in lower layers, leading to negative estimated flux values at the start of the
397 field sampling period. In this case it was only when the soil began to return to a baseline level
398 that the assumptions of the flux-gradient method were again met.

399 Both of the previous cases also provide context for the variable statistical comparisons between
400 field-measured soil fluxes and `neonSoilFlux` outputs (Figure 6). When considering systematic
401 deployment of this method across a measurement network, there are a number of independent
402 challenges that require careful consideration. There are clear tradeoffs between (1) accuracy
403 of modeled fluxes (defined here as closeness to field-measured F_S and the uncertainty reduc-
404 tion factor ϵ), (2) precision (which could be defined by the signal to noise ratio), and (3) the
405 choice of the diffusivity model (Section 4.2.2) or flux computation method (Section 4.2.3). A
406 sensitivity analysis (Figure S2, Supplemental Information) found that flux output uncertainty
407 was dominated by measurement uncertainty (T_S , P , SWC , or CO₂) rather than by the dif-
408 fusivity method used to compute soil flux. Notably, the F_{110} method was least sensitive to
409 measurement uncertainty likely because it best aligns with the surface chamber measurement
410 assumptions.

411 Finally, comparing the effects of different diffusivity estimation methods on the match between
412 modeled and measured fluxes (Figure 5) highlights the sensitivity of F_{ijk} to diffusivity. The
413 comparison between diffusivity estimates compared to field estimated diffusivity (Figure 7)

414 demonstrates that site parameters can dictate which measure of diffusivity is most likely to
415 be accurate in a given environmental context. Site-specific differences are largely a reflec-
416 tion of differences in soil moisture across the sites (Table 1), as not all diffusivity estimation
417 methods incorporate soil moisture equivalently. While we here have compares two approaches
418 to calculate diffusivity (the Millington-Quirk and Marshall models), it may be valuable to
419 evaluate other diffusivity models (e.g. the Moldrup model; Moldrup et al. (1999)) as well. Ul-
420 timately the choice of a particular diffusivity model could be determined based on knowledge
421 of site-specific evaluations or a set of these models could be used to generate a model ensemble
422 average as a means to trade precision for a more general approach.

423 **6.3 Recommendations for future method development**

424 The `neonSoilFlux` package provides several approaches to estimate soil flux using the gradient
425 method. We believe these approaches enable the software to be used across a range of site-
426 specific assumptions (Maier & Schack-Kirchner, 2014). We note, however, that this choice
427 can have a determinative approach on the calculated values. Ensemble averaging approaches
428 (Elshall et al., 2018; Raftery et al., 2005) may be one way to address this problem if the goal is
429 to calculate fluxes using the same method at a diverse range of different sites. Two other ideas
430 would be to apply machine learning algorithms (e.g. random forest) to generate a single flux
431 estimate across diverse sites, or using co-located estimates of net ecosystem carbon exchange
432 from eddy-flux towers to further constrain results or to assess soil flux results for plausibility
433 (Phillips et al., 2017).

434 These challenges notwithstanding, the method used here and made available in the
435 `neonSoilFlux` R package has the potential to produce nearly continuous estimates of flux
436 across all terrestrial NEON sites. These estimates are a significant improvement on available
437 approaches to constrain the portion of ecosystem respiration attributable to the soil. This,

⁴³⁸ in turn, also aids in our ability to understand the soil contribution to the net ecosystem flux
⁴³⁹ measured at these sites using the co-located eddy flux towers.

⁴⁴⁰ 7 Conclusions

⁴⁴¹ We used the R package `neonSoilFlux` to estimate soil CO₂ fluxes with the flux-gradient
⁴⁴² method using data from buried soil sensors at NEON terrestrial sites. We compared the
⁴⁴³ predicted fluxes to those measured directly using a field-based closed chamber approach. Soil
⁴⁴⁴ fluxes from `neonSoilFlux` were broadly effective at producing estimates of flux comparable
⁴⁴⁵ to those measured in the field using a chamber-based technique. However `neonSoilFlux`
⁴⁴⁶ outputs are quite sensitive to a number of issues, including: missing data (and thus gap-
⁴⁴⁷ filling of input measurement datasets), the selection of soil depths used to best calculate the
⁴⁴⁸ gradient (which may vary between sites), and finally the choice of method used for estimating
⁴⁴⁹ soil diffusivity. The flexibility of the `neonSoilFlux` package allows the user to evaluate each
⁴⁵⁰ of these issues with site-specific knowledge and contexts. Future refinements and subsequent
⁴⁵¹ validation of `neonSoilFlux` outputs will feed forward into evaluating soil carbon fluxes broader
⁴⁵² spatial scales to enhance understanding of the ways in which soils across diverse ecosystems
⁴⁵³ are responding to a changing climate.

⁴⁵⁴ Sources Cited

⁴⁵⁵ Ayres, E., Reichle, R. H., Colliander, A., Cosh, M. H., & Smith, L. (2024). Validation of
⁴⁵⁶ Remotely Sensed and Modeled Soil Moisture at Forested and Unforested NEON Sites.
⁴⁵⁷ *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17,
⁴⁵⁸ 14248–14264. <https://doi.org/10.1109/JSTARS.2024.3430928>

- 459 Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the
460 atmosphere - the state and future of the eddy covariance method. *Global Change Biology*,
461 20(12), 3600–3609. <https://doi.org/10.1111/gcb.12649>
- 462 Berenbaum, M. R., Carpenter, S. R., Hampton, S. E., Running, S. W., & Stanzione, D. C.
463 (2015). *Report from the NSF BIO Advisory Committee Subcommittee on NEON Scope
464 Impacts.*
- 465 Bond-Lamberty, B. (2018). New Techniques and Data for Understanding the Global Soil Res-
466piration Flux. *Earth's Future*, 6(9), 1176–1180. <https://doi.org/10.1029/2018EF000866>
- 467 Bond-Lamberty, B., Ballantyne, A., Berryman, E., Fluet-Chouinard, E., Jian, J., Morris, K.
468 A., Rey, A., & Vargas, R. (2024). Twenty Years of Progress, Challenges, and Opportuni-
469ties in Measuring and Understanding Soil Respiration. *Journal of Geophysical Research:
470 Biogeosciences*, 129(2), e2023JG007637. <https://doi.org/10.1029/2023JG007637>
- 471 Bond-Lamberty, B., Christianson, D. S., Malhotra, A., Pennington, S. C., Sihi, D., AghaK-
472 ouchak, A., Anjileli, H., Altaf Arain, M., Armesto, J. J., Ashraf, S., Ataka, M., Baldocchi,
473 D., Andrew Black, T., Buchmann, N., Carbone, M. S., Chang, S.-C., Crill, P., Curtis, P.
474 S., Davidson, E. A., ... Zou, J. (2020). COSORE: A community database for continuous
475 soil respiration and other soil-atmosphere greenhouse gas flux data. *Global Change Biology*,
476 26(12), 7268–7283. <https://doi.org/10.1111/gcb.15353>
- 477 Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Bio-
478 geosciences*, 7(6), 1915–1926. <https://doi.org/10.5194/bg-7-1915-2010>
- 479 Bond-Lamberty, B., Wang, C., & Gower, S. T. (2004). A global relationship between the
480 heterotrophic and autotrophic components of soil respiration? *Global Change Biology*,
481 10(10), 1756–1766. <https://doi.org/10.1111/j.1365-2486.2004.00816.x>
- 482 Bouma, T. J., & Bryla, D. R. (2000). On the assessment of root and soil respiration for soils
483 of different textures: Interactions with soil moisture contents and soil CO₂ concentrations.
484 *Plant and Soil*, 227(1), 215–221. <https://doi.org/10.1023/A:1026502414977>

- 485 Chen, H., & Tian, H.-Q. (2005). Does a General Temperature-Dependent Q10 Model of Soil
486 Respiration Exist at Biome and Global Scale? *Journal of Integrative Plant Biology*, 47(11),
487 1288–1302. <https://doi.org/10.1111/j.1744-7909.2005.00211.x>
- 488 Davidson, E. A., Janssens, I. A., & Luo, Y. (2006). On the variability of respiration in
489 terrestrial ecosystems: Moving beyond Q10. *Global Change Biology*, 12, 154–164. <https://doi.org/10.1111/j.1365-2486.2005.01065.x>
- 490 Desai, A. R., Murphy, B. A., Wiesner, S., Thom, J., Butterworth, B. J., Koupaei-Abyazani, N.,
491 Muttaqin, A., Paleri, S., Talib, A., Turner, J., Mineau, J., Merrelli, A., Stoy, P., & Davis,
492 K. (2022). Drivers of Decadal Carbon Fluxes Across Temperate Ecosystems. *Journal of
493 Geophysical Research: Biogeosciences*, 127(12), e2022JG007014. <https://doi.org/10.1029/2022JG007014>
- 494 Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman and
495 Hall/CRC. <https://doi.org/10.1201/9780429246593>
- 496 Elshall, A. S., Ye, M., Pei, Y., Zhang, F., Niu, G.-Y., & Barron-Gafford, G. A. (2018). Relative
497 model score: A scoring rule for evaluating ensemble simulations with application to micro-
498 bial soil respiration modeling. *Stochastic Environmental Research and Risk Assessment*,
499 32(10), 2809–2819. <https://doi.org/10.1007/s00477-018-1592-3>
- 500 Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G.,
501 Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger,
502 D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., ... Wofsy, S. (2001).
503 Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural
504 and Forest Meteorology*, 107(1), 43–69. [https://doi.org/10.1016/S0168-1923\(00\)00225-2](https://doi.org/10.1016/S0168-1923(00)00225-2)
- 505 Farrance, I., & Frenkel, R. (2012). Uncertainty of Measurement: A Review of the Rules
506 for Calculating Uncertainty Components through Functional Relationships. *The Clinical
507 Biochemist Reviews*, 33(2), 49–75.
- 508 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer,
509

- 511 P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz,
512 J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., ...
513 Zeng, J. (2025). Global Carbon Budget 2024. *Earth System Science Data*, 17(3), 965–1039.
514 <https://doi.org/10.5194/essd-17-965-2025>
- 515 Hamdi, S., Moyano, F., Sall, S., Bernoux, M., & Chevallier, T. (2013). Synthesis analysis
516 of the temperature sensitivity of soil respiration from laboratory studies in relation to
517 incubation methods and soil conditions. *Soil Biology and Biochemistry*, 58, 115–126. <https://doi.org/10.1016/j.soilbio.2012.11.012>
- 519 Hirano, T., Kim, H., & Tanaka, Y. (2003). Long-term half-hourly measurement of soil CO₂
520 concentration and soil respiration in a temperate deciduous forest. *Journal of Geophysical
521 Research: Atmospheres*, 108(D20). <https://doi.org/10.1029/2003JD003766>
- 522 Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017).
523 The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls.
524 *Annual Review of Ecology, Evolution and Systematics*, 48(Volume 48, 2017), 419–445.
525 <https://doi.org/10.1146/annurev-ecolsys-112414-054234>
- 526 Jian, J., Bailey, V., Dorheim, K., Konings, A. G., Hao, D., Shiklomanov, A. N., Snyder, A.,
527 Steele, M., Teramoto, M., Vargas, R., & Bond-Lamberty, B. (2022). Historically inconsis-
528 tent productivity and respiration fluxes in the global terrestrial carbon cycle. *Nature
529 Communications*, 13(1), 1733. <https://doi.org/10.1038/s41467-022-29391-5>
- 530 Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N.,
531 Manzon, J., Marchesi, R., Paredes, D., & Bond-Lamberty, B. (2021). A restructured and
532 updated global soil respiration database (SRDB-V5). *Earth System Science Data*, 13(2),
533 255–267. <https://doi.org/10.5194/essd-13-255-2021>
- 534 Jiang, J., Feng, L., Hu, J., Liu, H., Zhu, C., Chen, B., & Chen, T. (2024). Global soil
535 respiration predictions with associated uncertainties from different spatio-temporal data
536 subsets. *Ecological Informatics*, 82, 102777. <https://doi.org/10.1016/j.ecoinf.2024.102777>

- 537 Jobbágy, E. G., & Jackson, R. B. (2000). The Vertical Distribution of Soil Organic Carbon
538 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](https://doi.org/10.1890/1051-0761(2000)010%5B0423:TVDOSO%5D2.0.CO;2)
- 540 Liu, K., Li, X., Wang, S., & Zhang, H. (2023). A robust gap-filling approach for European
541 Space Agency Climate Change Initiative (ESA CCI) soil moisture integrating satellite
542 observations, model-driven knowledge, and spatiotemporal machine learning. *Hydrology*
543 and *Earth System Sciences*, 27(2), 577–598. <https://doi.org/10.5194/hess-27-577-2023>
- 544 Lunch, C., Laney, C., Mietkiewicz, N., Sokol, E., Cawley, K., & NEON (National Ecological
545 Observatory Network). (2025). *neonUtilities: Utilities for working with NEON data*. <https://doi.org/10.32614/CRAN.package.neonUtilities>
- 547 Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011).
548 Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*,
549 21(5), 1429–1442. <https://doi.org/10.1890/09-1275.1>
- 550 Maier, M., & Schack-Kirchner, H. (2014). Using the gradient method to determine soil gas
551 flux: A review. *Agricultural and Forest Meteorology*, 192–193, 78–95. <https://doi.org/10.1016/j.agrformet.2014.03.006>
- 553 Mariethoz, G., Linde, N., Jougnot, D., & Rezaee, H. (2015). Feature-preserving interpolation
554 and filtering of environmental time series. *Environmental Modelling & Software*, 72, 71–76.
555 <https://doi.org/10.1016/j.envsoft.2015.07.001>
- 556 Marshall, T. J. (1959). The Diffusion of Gases Through Porous Media. *Journal of Soil Science*,
557 10(1), 79–82. <https://doi.org/10.1111/j.1365-2389.1959.tb00667.x>
- 558 Millington, R. J., & Shearer, R. C. (1971). Diffusion in aggregated porous media. *Soil Science*,
559 111(6), 372–378.
- 560 Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G.,
561 Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann,
562 M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., & Stauch, V. J. (2007). Comprehensive

- 563 comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and*
564 *Forest Meteorology*, 147(3), 209–232. <https://doi.org/10.1016/j.agrformet.2007.08.011>
- 565 Moldrup, P., Olesen, T., Yamaguchi, T., Schjønning, P., & Rolston, D. E. (1999). Modeling
566 diffusion and reaction in soils: 9. The Buckingham-Burdine-Campbell equation for gas
567 diffusivity in undisturbed soil. *Soil Science*, 164(2), 75.
- 568 National Ecological Observatory Network (NEON). (2024a). *Barometric pressure*
569 (*DP1.00004.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/RT4V-KZ04>
- 570 National Ecological Observatory Network (NEON). (2024b). *Soil CO₂ concentration*
571 (*DP1.00095.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/E7GR-6G94>
- 572 National Ecological Observatory Network (NEON). (2024c). *Soil physical and chemical properties*,
573 *Megapit* (*DP1.00096.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/S6ND-Q840>
- 574 National Ecological Observatory Network (NEON). (2024d). *Soil temperature* (*DP1.00041.001*).
575 National Ecological Observatory Network (NEON). <https://doi.org/10.48443/Q24X-PW21>
- 576 National Ecological Observatory Network (NEON). (2024e). *Soil water content and water*
577 *salinity* (*DP1.00094.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/A8VY-Y813>
- 578 Norman, J. M., Kucharik, C. J., Gower, S. T., Baldocchi, D. D., Crill, P. M., Rayment, M.,
579 Savage, K., & Striegl, R. G. (1997). A comparison of six methods for measuring soil-
580 surface carbon dioxide fluxes. *Journal of Geophysical Research: Atmospheres*, 102(D24),
581 28771–28777. <https://doi.org/10.1029/97JD01440>
- 582 Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown,
583 K., & Vargas, R. (2017). The value of soil respiration measurements for interpreting and
584 modeling terrestrial carbon cycling. *Plant and Soil*, 413(1), 1–25. <https://doi.org/10.1007/>

- 589 s11104-016-3084-x
- 590 Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). *Using Bayesian Model*
591 *Averaging to Calibrate Forecast Ensembles.* <https://doi.org/10.1175/MWR2906.1>
- 592 Sallam, A., Jury, W. A., & Letey, J. (1984). Measurement of Gas Diffusion Coefficient under
593 Relatively Low Air-filled Porosity. *Soil Science Society of America Journal*, 48(1), 3–6.
594 <https://doi.org/10.2136/sssaj1984.03615995004800010001x>
- 595 Shao, J., Zhou, X., Luo, Y., Li, B., Aurela, M., Billesbach, D., Blanken, P. D., Bracho, R.,
596 Chen, J., Fischer, M., Fu, Y., Gu, L., Han, S., He, Y., Kolb, T., Li, Y., Nagy, Z., Niu, S.,
597 Oechel, W. C., ... Zhang, J. (2015). Biotic and climatic controls on interannual variability
598 in carbon fluxes across terrestrial ecosystems. *Agricultural and Forest Meteorology*, 205,
599 11–22. <https://doi.org/10.1016/j.agrformet.2015.02.007>
- 600 Shao, P., Zeng, X., Moore, D. J. P., & Zeng, X. (2013). Soil microbial respiration from
601 observations and Earth System Models. *Environmental Research Letters*, 8(3), 034034.
602 <https://doi.org/10.1088/1748-9326/8/3/034034>
- 603 Sih, D., Gerber, S., Inglett, P. W., & Inglett, K. S. (2016). Comparing models of microbial–
604 substrate interactions and their response to warming. *Biogeosciences*, 13(6), 1733–1752.
605 <https://doi.org/10.5194/bg-13-1733-2016>
- 606 Tang, J., Baldocchi, D. D., Qi, Y., & Xu, L. (2003). Assessing soil CO₂ efflux using continuous
607 measurements of CO₂ profiles in soils with small solid-state sensors. *Agricultural and Forest*
608 *Meteorology*, 118(3), 207–220. [https://doi.org/10.1016/S0168-1923\(03\)00112-6](https://doi.org/10.1016/S0168-1923(03)00112-6)
- 609 Tang, J., Misson, L., Gershenson, A., Cheng, W., & Goldstein, A. H. (2005). Continuous
610 measurements of soil respiration with and without roots in a ponderosa pine plantation
611 in the Sierra Nevada Mountains. *Agricultural and Forest Meteorology*, 132(3), 212–227.
612 <https://doi.org/10.1016/j.agrformet.2005.07.011>
- 613 Taylor, J. R. (2022). *An Introduction to Error Analysis: The Study of Uncertainties in Physical*
614 *Measurements, Third Edition* (3rd ed.). University Science Press.

- 615 Yan, Z., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S., Liu, C., & Liu, C. (2018).
616 A moisture function of soil heterotrophic respiration that incorporates microscale processes.
617 *Nature Communications*, 9(1), 2562. <https://doi.org/10.1038/s41467-018-04971-6>
- 618 Yan, Z., Liu, C., Todd-Brown, K. E., Liu, Y., Bond-Lamberty, B., & Bailey, V. L. (2016).
619 Pore-scale investigation on the response of heterotrophic respiration to moisture conditions
620 in heterogeneous soils. *Biogeochemistry*, 131(1), 121–134. <https://doi.org/10.1007/s10533-016-0270-0>
- 622 Zhang, R., Kim, S., Kim, H., Fang, B., Sharma, A., & Lakshmi, V. (2023). Temporal
623 Gap-Filling of 12-Hourly SMAP Soil Moisture Over the CONUS Using Water Balance
624 Budgeting. *Water Resources Research*, 59(12), e2023WR034457. <https://doi.org/10.1029/2023WR034457>
- 626 Zobitz, J., Ayres, E., O'Rourke, K., Werbin, Z., Lee, L., Abdi, R., Mehmeti, D., & Xiong, L.
627 (2024). *neonSoilFlux: Compute Soil Carbon Fluxes for the National Ecological Observatory*
628 *Network Sites*. <https://doi.org/10.32614/CRAN.package.neonSoilFlux>