

<sup>1</sup> **neonSoilFlux: An R Package for Continuous  
2 Sensor-Based Estimation of Soil CO<sub>2</sub> Fluxes**

<sup>3</sup> John Zobitz<sup>1</sup>      Edward Ayres<sup>2</sup>      Zoey Werbin<sup>3</sup>      Ridwan Abdi<sup>1</sup>  
<sup>4</sup>                        Natalie Ashburner-Wright<sup>4</sup>      Lillian Brown<sup>4</sup>  
<sup>5</sup>                        Ryan Frink-Sobierajski<sup>4</sup>      Lajntxiag Lee<sup>1</sup>      Dijonë Mehmeti<sup>1</sup>  
<sup>6</sup>                        Christina Tran<sup>4</sup>      Ly Xiong<sup>1</sup>      Naupaka Zimmerman<sup>4,5</sup>

<sup>7</sup> <sup>1</sup> Augsburg University, 2211 Riverside Avenue, Minneapolis, MN 55454

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

<sup>10</sup> <sup>3</sup> Boston University, 5 Cummington Street, Boston, MA 02215

<sup>11</sup> <sup>4</sup> University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117

<sup>12</sup> <sup>5</sup> University of Kansas, 1450 Jayhawk Boulevard, Lawrence, KS 66045

<sup>13</sup> **Acknowledgments**

<sup>14</sup> John Zobitz acknowledges Kathleen O'Rourke for code development. Naupaka Zimmerman  
<sup>15</sup> thanks technical staff at USF for support with field gear assembly and shipping. We thank the  
<sup>16</sup> NEON field staff and assignable assets teams for facilitating each of the six NEON site visits.

<sup>17</sup> We are grateful to LI-COR technical staff for helpful discussions about optimal soil chamber  
<sup>18</sup> sampling methods. This work was supported by NSF DEB grant #2017829 awarded to John  
<sup>19</sup> Zobitz, and NSF DEB grant #2017860 awarded to Naupaka Zimmerman. This material is  
<sup>20</sup> based in part upon work supported by the National Ecological Observatory Network (NEON),  
<sup>21</sup> a program sponsored by the U.S. National Science Foundation (NSF) and operated under  
<sup>22</sup> cooperative agreement by Battelle. We also thank the reviewers and subject editor for their  
<sup>23</sup> constructive feedback.

## <sup>24</sup> **Conflict of Interest Statements**

<sup>25</sup> None of the authors have a financial, personal, or professional conflict of interest related to  
<sup>26</sup> this work.

## <sup>27</sup> **Author Contributions**

<sup>28</sup> Conceptualization: John Zobitz, Naupaka Zimmerman; Methodology: Edward Ayres, John  
<sup>29</sup> Zobitz, Naupaka Zimmerman; Software: John Zobitz, Naupaka Zimmerman, Zoey Werbin,  
<sup>30</sup> Edward Ayres, Dijonë Mehmeti, Ridwan Abdi, Ly Xiong, Lajntxiag Lee; Validation: John  
<sup>31</sup> Zobitz, Naupaka Zimmerman; Formal Analysis: John Zobitz, Naupaka Zimmerman, Dijonë  
<sup>32</sup> Mehmeti, Ridwan Abdi, Ly Xiong, Lajntxiag Lee; Investigation: John Zobitz, Naupaka Zimmerman,  
<sup>33</sup> Ryan Frink-Sobierajski, Christina Tran, Natalie Ashburner-Wright, Lillian Brown;  
<sup>34</sup> Resources: John Zobitz, Naupaka Zimmerman; Data curation: John Zobitz, Naupaka Zimmerman,  
<sup>35</sup> Dijonë Mehmeti, Ly Xiong; Writing – original draft: John Zobitz, Naupaka Zimmerman;  
<sup>36</sup> Writing – review and editing: John Zobitz, Naupaka Zimmerman, Zoey Werbin, Edward Ayres,  
<sup>37</sup> Christina Tran, Dijonë Mehmeti, Ly Xiong; Visualization: John Zobitz, Naupaka Zimmerman,

<sup>38</sup> Dijonë Mehmeti, Ridwan Abdi, Ly Xiong; Supervision: John Zobitz, Naupaka Zimmerman;  
<sup>39</sup> Project Administration: John Zobitz, Naupaka Zimmerman; Funding Acquisition: John Zob-  
<sup>40</sup> itz, Naupaka Zimmerman.

<sup>41</sup> **Data Availability**

<sup>42</sup> Anonymous field-collected data, `neonSoilFlux` calculated outputs, and manuscript-generating  
<sup>43</sup> code for peer review are provided as supplemental files. An anonymous link for peer-review  
<sup>44</sup> is here: <https://doi.org/10.5281/zenodo.1695117>. This will be made publicly available upon  
<sup>45</sup> publication.

46 **1 Abstract**

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

<sup>69</sup> **1.1 Keywords**

<sup>70</sup> Soil carbon, carbon dioxide, flux gradient, carbon cycle, field validation, soil respiration, ecosys-  
<sup>71</sup> tem variability, diffusion

<sup>72</sup> **2 Data for peer review**

<sup>73</sup> Anonymous field-collected data, `neonSoilFlux` calculated outputs, and manuscript-generating  
<sup>74</sup> code for peer review are provided as supplemental files. An anonymous link for peer-review  
<sup>75</sup> is here: <https://doi.org/10.5281/zenodo.1695117>. This will be made publicly available upon  
<sup>76</sup> publication.

<sup>77</sup> **3 Introduction**

<sup>78</sup> Soils contain the planet's largest reservoir of terrestrial carbon (Jobbágy & Jackson, 2000). A  
<sup>79</sup> critical component of this reservoir is soil organic matter, the accumulation of which is influ-  
<sup>80</sup> enced by biotic factors such as above-ground plant inputs (Jackson et al., 2017). These inputs  
<sup>81</sup> in turn are influenced by environmental factors such as growing season length, temperature,  
<sup>82</sup> and moisture (Desai et al., 2022), which also affect the breakdown of soil organic matter and its  
<sup>83</sup> return to the atmosphere. Across heterogeneous terrestrial landscapes, the interplay between  
<sup>84</sup> these biotic and abiotic factors influence the size of the soil contribution to the terrestrial  
<sup>85</sup> carbon sink (Friedlingstein et al., 2025). However, the heterogeneity of these processes across  
<sup>86</sup> diverse ecosystems in the context of rapid environmental change leads to large uncertainty  
<sup>87</sup> about the magnitude of this sink in the future, and thus there remains a pressing need to  
<sup>88</sup> quantify changes in soil carbon pools and fluxes across scales.

89 Ecological observation networks such as the United States' National Ecological Observatory  
90 Network (NEON) and others (e.g. the globally-distributed FLUXNET or the European Inte-  
91 grated Carbon Observation System) present a significant advancement in the nearly continuous  
92 observation of biogeochemical processes at the continental scale. Notably, at 47 terrestrial sites  
93 across the continental United States that span 20 ecoclimatic domains, NEON provides half-  
94 hourly measurements of soil CO<sub>2</sub> concentration, temperature, and moisture at different vertical  
95 depths. Each of these NEON sites also encompasses measurements of the cumulative sum of all  
96 ecosystem carbon fluxes in an airshed using the eddy covariance technique (Baldocchi, 2014).  
97 Soil observations provided by NEON are on the same timescale and standardized with eddy co-  
98 variance measurements from FLUXNET. These types of nearly continuous observational data  
99 (NEON and FLUXNET) can be used to reconcile differences between model-derived or data-  
100 estimated components of ecosystem carbon flux (Jian et al., 2022; Luo et al., 2011; Phillips et  
101 al., 2017; J. Shao et al., 2015; P. Shao et al., 2013; Sihl et al., 2016).

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

109 One common method by which  $F_S$  is measured in the field is through the use of soil chambers  
110 in a closed, well-mixed system (Norman et al., 1997) with headspace trace gas concentrations  
111 measured with an infrared gas analyzer (IRGA).  $F_S$  can also be estimated from soil CO<sub>2</sub>  
112 measurements at different depths in the soil using the flux-gradient method (Maier & Schack-  
113 Kirchner, 2014). Closed-chamber IRGA measurements, while being the most common method,

require either frequent in-person site visits or expensive and fragile automated systems. The potential of the gradient method is that fluxes can be estimated from continuous data recorded by robust solid-state sensors. The flux-gradient method is an approach that uses conservation of mass to calculate flux at a 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 is no mass transfer in the other spatial dimensions  $x$  and  $y$  (Maier & Schack-Kirchner, 2014). The diffusivity profile, a key component of this calculation, varies across the soil depth as a function of soil temperature, soil volumetric water content, atmospheric air pressure, and soil bulk density (Millington & Shearer, 1971; Moldrup et al., 1999; Sallam et al., 1984).

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

This study describes an R software package, `neonSoilFlux`, that computes a standardized estimate of  $F_S$  at all terrestrial NEON sites using the flux-gradient method. Using direct chamber-based field observations of soil carbon dioxide flux from a subset of terrestrial NEON sites spanning six states, we provide a direct validation of  $F_S$  from `neonSoilFlux`. While open source R software tools currently exist for processing chamber-based flux measurements (Jurasinski et al., 2022; Pedersen, 2024; Rheault et al., 2024; Wilson et al., 2024; Zhao, 2019),

<sup>139</sup> to our knowledge this is the first package that incorporates NEON data directly.

<sup>140</sup> Key objectives of this study are to:

- <sup>141</sup> 1. Apply the flux-gradient method to estimate soil CO<sub>2</sub> flux from continuous sensor mea-
- <sup>142</sup> surements across six NEON sites.
- <sup>143</sup> 2. Benchmark estimated soil carbon fluxes against field measurements (e.g. direct chamber
- <sup>144</sup> measurements of soil flux).
- <sup>145</sup> 3. Identify sources of error in the flux-gradient approach across diverse sites in order to
- <sup>146</sup> guide future work.

## <sup>147</sup> **4 Materials and Methods**

### <sup>148</sup> **4.1 Field methods**

#### <sup>149</sup> **4.1.1 Focal NEON Sites**

<sup>150</sup> In order to acquire field data to validate model predictions of flux, we selected six terrestrial  
<sup>151</sup> NEON sites for analysis. We conducted roughly week-long field measurement campaigns at  
<sup>152</sup> these sites, which span a range of environmental gradients and terrestrial domains (Table 1).  
<sup>153</sup> SJER, SRER, and WREF were visited during May and June of 2022, and WOOD, KONZ,  
<sup>154</sup> and UNDE during May and June of 2024.

#### <sup>155</sup> **4.1.2 Soil collar placement**

<sup>156</sup> Either one (2022 sampling campaign) or two (2024 sampling campaign) PVC soil collars (20.1  
<sup>157</sup> cm inside diameter) were installed in close proximity to the permanent NEON soil sensors at

158 each site (Figure 1). As instruments in the NEON soil sensor arrays can occasionally break  
159 down or stop working, the specific soil plot where we made measurements was chosen at each  
160 site in consultation with NEON staff to maximize likelihood of quality soil sensor measurements  
161 during the duration of the IRGA measurements. The plot selected at each site (out of the 5 in  
162 each replicate array at each site) are presented in the last column of Table 1. After installation,  
163 collar(s) were left to equilibrate for approximately 24 hours prior to any measurements being  
164 taken.

165 **4.1.3 Infrared gas analyzer measurements of soil CO<sub>2</sub> flux**

166 In 2022, we then made measurements of flux on an hourly interval for 8 hours each day.  
167 Measurements were taken from roughly 8 am to 4 pm, with the time interval selected to  
168 capture the majority of the diurnal gradient of soil temperature each day. These measurements  
169 were made using a LI-6800 infrared gas analyzer instrument (LI-COR Environmental, Lincoln,  
170 NE) fitted with a soil chamber attachment (attachment 6800-09). In 2024, we again used the  
171 same LI-6800 instrument, but made half-hourly measurements over an approximately 8 hour  
172 period. In addition, in 2024 we also installed a second collar and used a second instrument, an  
173 LI-870 CO<sub>2</sub> IRGA, connected to an automated robotic chamber (LI-COR chamber 8200-104)  
174 controlled by an LI-8250 multiplexer to make automated measurements. The multiplexer was  
175 configured to take half-hourly measurements 24 hours a day for the duration of our sampling  
176 bout at each site. Each instrument was paired with a soil temperature and moisture probe  
177 (Stevens HydraProbe, Stevens Water, Portland, OR) that was used to make soil temperature  
178 and moisture measurements concurrent with the CO<sub>2</sub> flux measurements. Chamber volumes  
179 were set by measuring collar offsets at each site. System checks were conducted daily for the  
180 LI-6800 and weekly for the LI-8250. Instruments were factory calibrated before each field  
181 season.

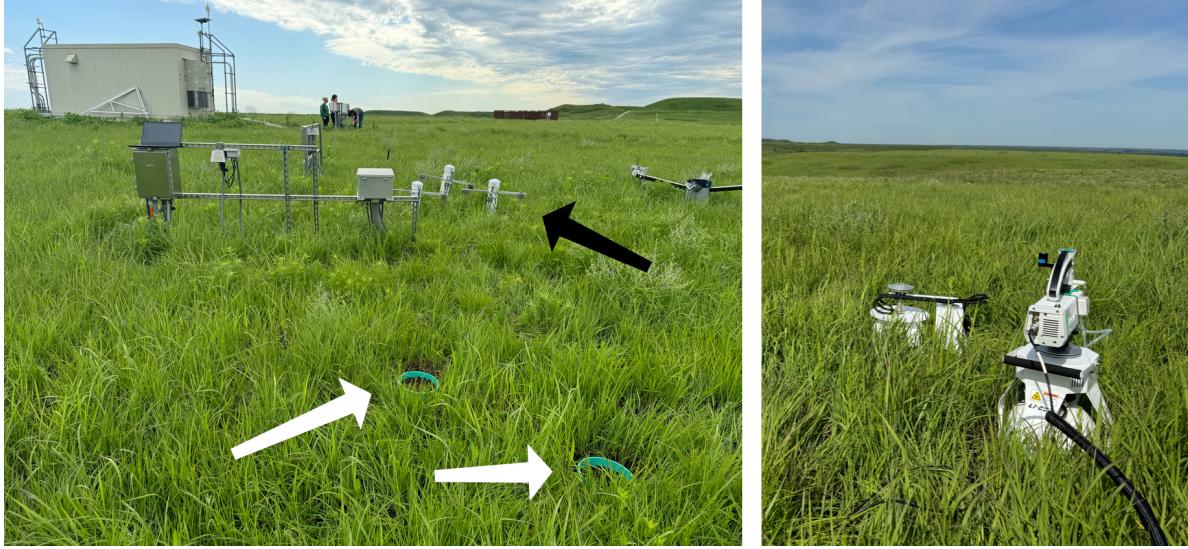


Figure 1: Spatial layout of field sampling using a closed-dynamic chamber setup at a representative NEON site (KONZ). Left image shows collars (white arrows) and permanent soil sensor installation (black arrow) and right image shows the LI-6800 (foreground) and LI-8200-104 (background) instruments placed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

#### 182 4.1.4 Post-collection processing of field data

183 We used LI-COR SoilFluxPro software (v 5.3.1) to assess the data after collection and to inform  
 184 sampling parameters. We checked appropriateness of dead band and measurement durations  
 185 using built-in evaluation tools. Based on this, the deadband period was set for 30-40 seconds,  
 186 depending on the site, and the measurement duration was 180 seconds with a 30 second pre-  
 187 purge and a 30 second post-purge at most sites, and a 90 second pre- and post-purge at sites  
 188 with higher humidity due to recent precipitation events. We also assessed the  $R^2$  of linear and  
 189 exponential model fits to measured CO<sub>2</sub> to verify measurement quality.

#### 190 4.2 neonSoilFlux R package

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

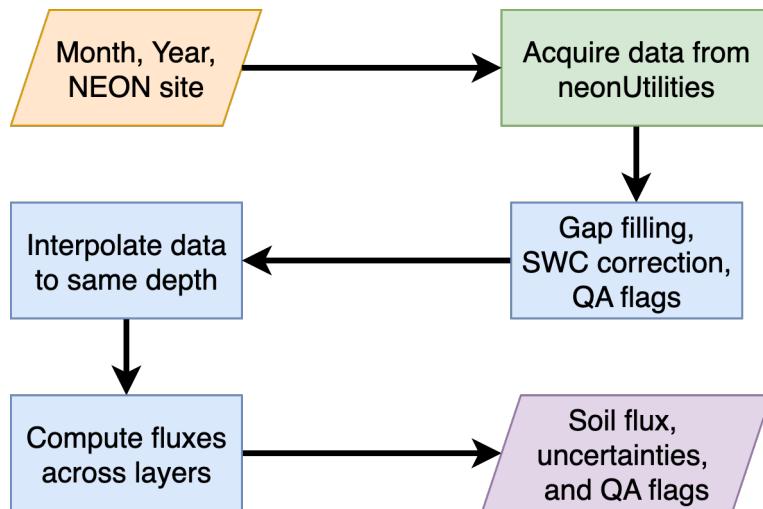


Figure 2: Diagram of `neonSoilFlux` R package. For a given month, year, and NEON site (orange parallelogram), the package acquires all relevant data to compute  $F_S$  using the `neonUtilities` R package (green rectangle). Data are gap-filled according to reported QA flags and adjusted for changes in 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).

At a given NEON site there are five replicate soil plots, each with measurements of soil  $\text{CO}_2$  concentration, soil temperature, and soil moisture at different depths (Figure 3). The `neonSoilFlux` package acquires measured soil  $\text{CO}_2$  concentration (NEON, 2024b), soil temperature (NEON, 2024d), soil water content (NEON, 2024e), barometric pressure from the nearby tower (NEON, 2024a), and soil properties (e.g. bulk density) (NEON, 2024c) from a range of different NEON data products. The static soil properties were collected by NEON staff from a nearby soil pit during initial site characterization and are assumed to be constant at each site. A soil flux calculation is computed at each replicate soil plot.

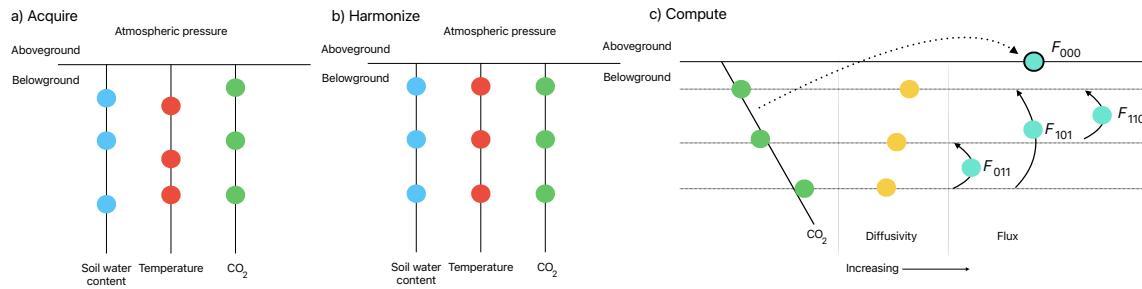


Figure 3: Model diagram of the data workflow for the `neonSoilFlux` R package. a) Acquire: Data are obtained for a given NEON location and horizontal sensor location, which includes soil water content, soil temperature,  $\text{CO}_2$  concentration, and atmospheric pressure. All data are screened for 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  $\text{CO}_2$  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 ( $i = \text{closest to surface}$ ,  $k = \text{deepest}$ ).  $F_{000}$  represents a flux estimate where the gradient  $d\text{C}/dz$  is the slope of a linear regression of  $\text{CO}_2$  with depth.

The workflow to compute a value of  $F_S$  with `neonSoilFlux` consists of three primary steps, illustrated in Figure 3. First, NEON data are acquired for a given site and month via the `neonUtilities` R package (yellow parallelogram and green rectangle in Figure 2 and Panel a in Figure 3). Acquired environmental data can be exported to a comma separated value file for additional analysis. Quality assurance (QA) flags are reported as an indicator variable. Since

208 the calibration coefficients on the soil water content sensors have changed over time (NEON,  
209 2024e), raw sensor measurements were back-calculated and soil-specific calibrations were ap-  
210 plied following Ayres et al. (2024) to generate a consistent time series at each measurement  
211 location.

212 The second step is harmonizing the data to compute soil fluxes across soil layers. This step  
213 consists of three different actions (blue rectangles in Figure 2 and Panel b in Figure 3). If a  
214 given observation by NEON is reported as not passing a quality assurance check, we applied  
215 a gap filling method to replace that measurement with its monthly mean at that same depth  
216 (Section 4.2.1). Belowground measurements of soil water and soil temperature are then inter-  
217 polated to the same depth as soil CO<sub>2</sub> measurements. The diffusivity (Section 4.2.2) and soil  
218 flux across different soil layers (Section 4.2.3) are then computed.

219 The third and final step is computing a surface soil flux through extrapolation to the sur-  
220 face (purple parallelogram in Figure 2 and Panel c in Figure 3). Uncertainty on a soil flux  
221 measurement is computed through quadrature. An aggregate quality assurance (QA) flag  
222 for each environmental measurement is also reported, representing if any gap-filled measure-  
223 ments were used in the computation of a soil flux. Within the soil flux-gradient method,  
224 several different approaches can be used to derive a surface flux (Maier & Schack-Kirchner,  
225 2014); the `neonSoilFlux` package reports four different possible values for soil surface flux  
226 (Section 4.2.3) for each of two different methods of diffusivity estimation, for a total of eight  
227 estimates of flux.

#### 228 **4.2.1 Gap-filling routine**

229 NEON reports QA flags as binary values for each measurement and half-hourly interval. For  
230 a given half-hour, if any input variable (soil CO<sub>2</sub> concentration, soil temperature, or soil

moisture) at depth  $z$  is flagged, computation of  $F_S$  is not possible. To address this, flagged measurements and their uncertainties were replaced with a bootstrapped monthly mean ( $\bar{m}$ ) and monthly standard deviation ( $\bar{s}$ ) (Efron & Tibshirani, 1994).

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

From these, 5000 bootstrap samples were generated for  $\mathbf{m}, \sigma$ , and  $\mathbf{b}$ . For each sample  $(m_k, b_k, \sigma_k)$ , we generated a vector  $\mathbf{n}$  (length  $N = 5000$ ) by drawing from a normal distribution with mean  $m_k + b_k$  and standard deviation  $\sigma_k$ . The sample mean and standard deviation were then computed from  $\mathbf{n}$ . The resulting distributions of sample means and sample standard deviations provided the bootstrapped monthly mean ( $\bar{m}$ ) and standard error ( $\bar{s}$ ) respectively.

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

#### 4.2.2 Soil diffusivity

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

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

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

253 where  $T_i$  is soil temperature ( $^{\circ}\text{C}$ ) at depth  $i$  (NEON, 2024d) and  $P$  surface barometric pressure  
 254 (kPa) (NEON, 2024a).

255 Previous studies by Sallam et al. (1984) and Tang et al. (2003) demonstrated the sensitivity  
 256 of modeled  $F_S$  depending on the tortuosity model ( $\xi$ ) used to compute diffusivity. At low  
 257 soil water content, the choice of tortuosity model can lead to order-of-magnitude differences  
 258 in  $D_a$ , which in turn affect modeled  $F_S$ . The `neonSoilFlux` package currently includes two  
 259 approaches to calculate  $\xi$ , representing the range of tortuosity behavior reported in Sallam et  
 260 al. (1984).

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

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

263 In Equation 2,  $SWC$  is the soil water content at depth  $i$  (NEON, 2024e) and  $\phi$  is the porosity,  
 264 which in turn is a function of soil physical properties (NEON, 2024c):

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

265 In Equation 3,  $\rho_m$  is the particle density of mineral soil ( $2.65 \text{ g cm}^{-3}$ ),  $\rho_s$  the soil bulk density ( $\text{g}$   
 266  $\text{cm}^{-3}$ ) excluding coarse fragments greater than 2 mm (NEON, 2024c), and  $f_V$  is a site-specific  
 267 value that accounts for the proportion of soil fragments between 2-20 mm. Soil fragments

268 greater than 20 mm were not estimated due to limitations in the amount of soil that can be  
269 analyzed (NEON, 2024c). We assume that rock fragments contain no internal pores.

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

276 **4.2.3 Soil flux computation**

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

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

279 where  $D_a$  is the diffusivity ( $\text{m}^2 \text{s}^{-1}$ ) and  $\frac{dC}{dz}$  is the gradient of CO<sub>2</sub> molar concentration  
280 ( $\mu\text{mol m}^{-3}$ , so the gradient has units of  $\mu\text{mol m}^{-3} \text{m}^{-1}$ ). The soil surface flux is theoretically  
281 defined by applying Equation 4 to measurements collected at the soil surface and directly  
282 below the surface. Measurements of soil temperature, soil water content, and soil CO<sub>2</sub> molar  
283 concentration across the soil profile allow for application of Equation 4 across different soil  
284 depths. Each site had three measurement layers, so we denote the flux as a three-digit subscript  
285  $F_{ijk}$  with indicator variables  $i$ ,  $j$ , and  $k$  indicate if a given layer was used (written in order of  
286 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<sub>2</sub> 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.
- $F_{110}$  is a flux estimate across the two shallowest measurement layers.
- $F_{011}$  is a flux estimate across the two deepest measurement layers.
- $F_{101}$  is a flux estimate across the shallowest and deepest measurement layers.

For  $F_{110}$ ,  $F_{011}$ , and  $F_{101}$ , the diffusivity used in Fick's Law is always at the deeper measurement layer. When used as a surface flux estimate we assume CO<sub>2</sub> remains constant above this flux depth. Uncertainty in all  $F_{ijk}$  values was quantified using quadrature (Taylor, 2022). These computed fluxes could provide the basis for additional soil flux estimates. For example, Tang et al. (2005) estimated surface flux by linearly extrapolating  $F_{110}$  and  $F_{011}$  to the soil surface.

### 4.3 Post processing evaluation

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

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 CO<sub>2</sub> 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 <sup>-2</sup> s <sup>-1</sup>	Soil temp °C	VSWC cm <sup>3</sup> cm <sup>-3</sup>	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

## 5 Results

### Concordance between modelled and measured soil CO<sub>2</sub> flux

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

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

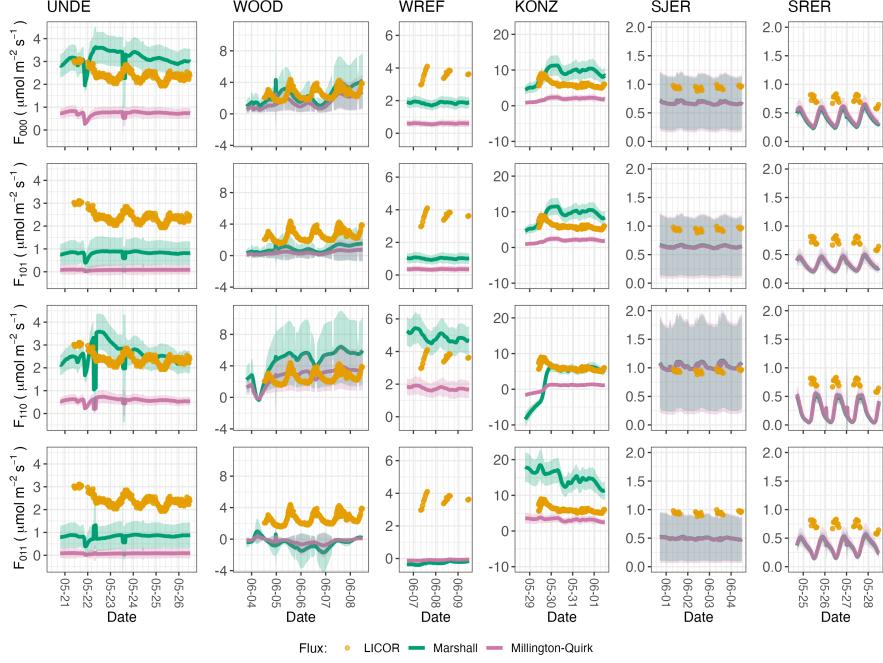


Figure 4: Timeseries of soil surface flux ( $F_S$ ) from field-measured (yellow lines) and modeled soil fluxes (green or purple lines) by the `neonSoilFlux` R package. Fluxes from the `neonSoilFlux` R package are separated by the diffusivity model used (Millington-Quirk or Marshall, Section 4.2.2). Individual 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  $\pm 1$  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.

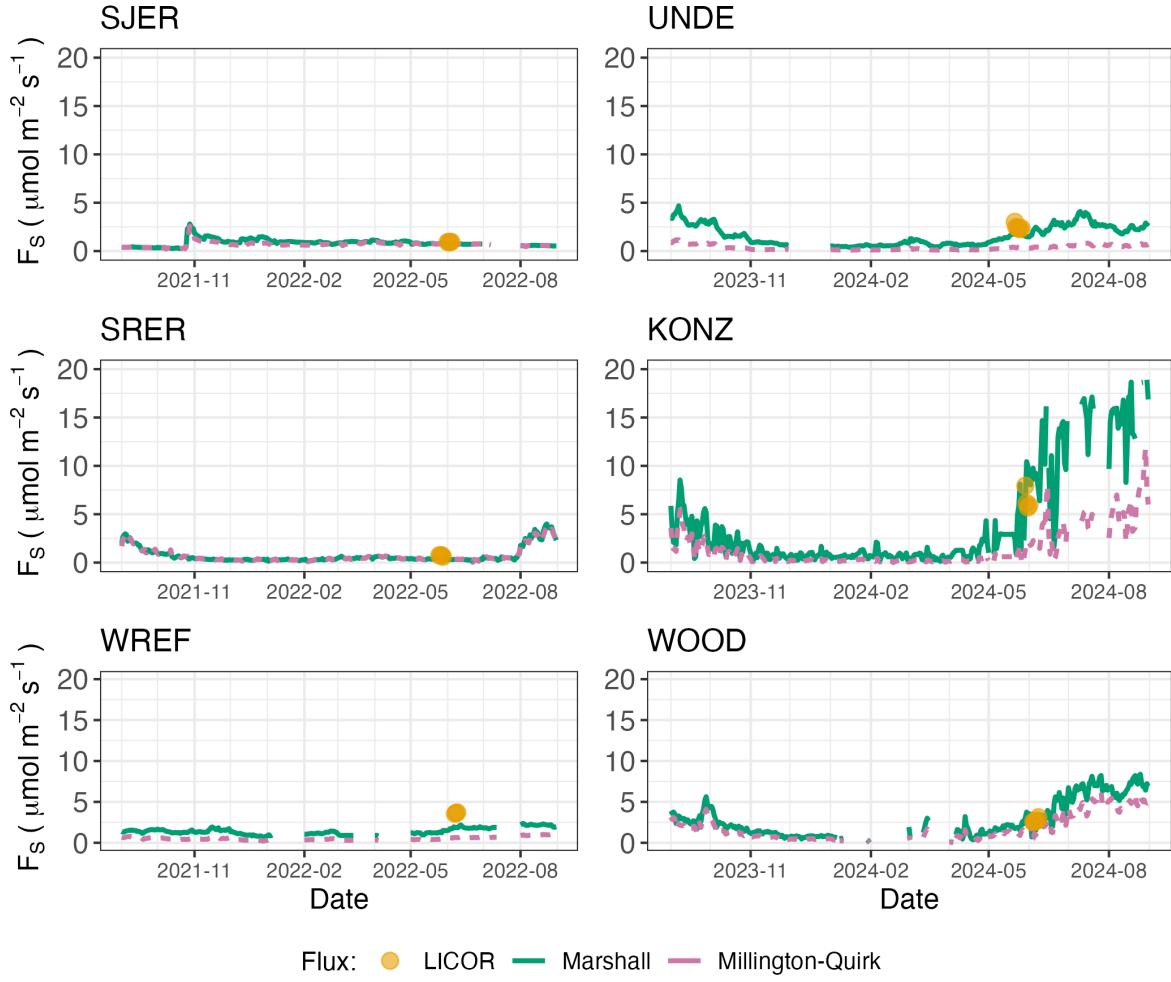


Figure 5: Timeseries of both daily-averaged field  $F_S$  (yellow circles) and daily ensemble averaged soil fluxes (average of  $F_{000}$ ,  $F_{101}$ ,  $F_{011}$ ,  $F_{110}$ , Section 4.2.3) by the `neonSoilFlux` R package, separated by the diffusivity model used (green or purple lines, Millington-Quirk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

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

324 **5.2 Effects of method choice on diffusivity estimates**

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

329 **6 Discussion**

330 This study presents a unified data science workflow to efficiently process automated measure-  
331 ments of belowground soil CO<sub>2</sub> concentrations, soil water content, and soil temperature to  
332 infer estimates of soil surface CO<sub>2</sub> effluxes through application of Fick's Law (Equation 4).  
333 Our core goals in this study were: (1) to generate estimates of soil flux from continuous soil  
334 sensor data at terrestrial NEON sites using the flux-gradient method and then (2) to compare  
335 those estimates to field-measured fluxes based on the closed chamber approach at six NEON  
336 focal sites. We discuss our progress toward these core goals through (1) an overall evaluation  
337 of the flux-gradient approach (and uncertainty calculation) and (2) site-specific evaluation of  
338 differences in estimated vs measured fluxes.

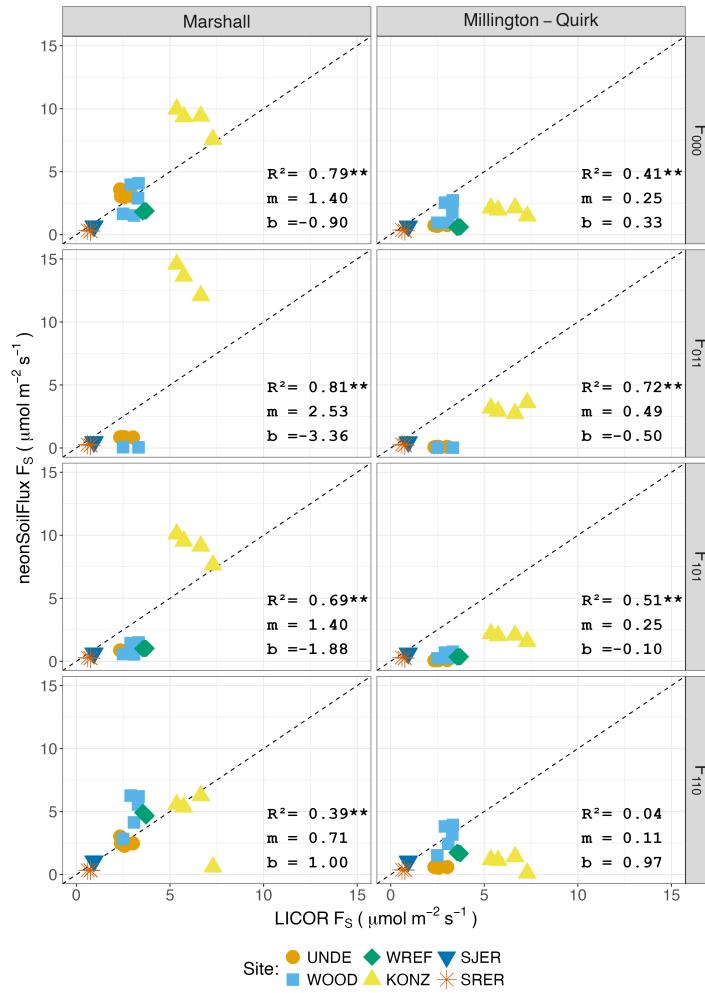


Figure 6: Statistical comparison between measured fluxes at each NEON site with fluxes reported by `neonSoilFlux` with the different flux calculation approaches and diffusivity calculations applied. Points are daily averages and LICOR  $F_S$  values are from the 6800 instrument only, for consistency. 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 slope ( $m$ ) and intercept ( $b$ ) of the linear regression between measured and modeled fluxes are also reported. The low-value outlier from KONZ in the  $F_{110}$  Marshall plot is an example of the effect of inverted CO<sub>2</sub> gradients causing an estimated flux to be negative, bringing down the daily mean, which later resolved as the soils dried back out.

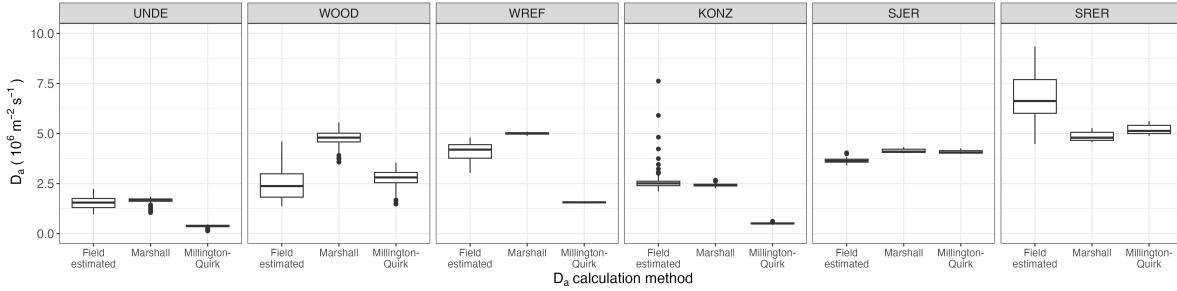


Figure 7: Distribution of diffusivity ( $D_a$ ) at each study site. Values of  $D_a$  were provided by the `neonSoilFlux` package, computed from the Millington-Quirk or Marshall models (Section 4.2.2). A post-hoc estimate of diffusivity (labeled as “Field estimated”) was computed through the field measured flux (Figure 4), divided by the  $\text{CO}_2$  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.

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

Key assumptions of the flux-gradient approach are that  $\text{CO}_2$  concentrations increase throughout the soil profile such that the highest concentrations are observed in the deepest layers. Additionally, field flux measurements should correlate with  $F_{000}$  because they represent surface fluxes. Periods where this gradient condition are not met generally are connected to processes that occur during soil wetting events, where more shallow soil layers produce higher concentrations of  $\text{CO}_2$  due to microbial respiration pulses following rewetting. This effect is likely to be largest at sites with rich organic soils (e.g. KONZ). Based on this reasoning, in these types of situations we would *a priori* expect  $F_{011}$  (deepest layers)  $\leq F_{101} \leq F_{110}$  (shallow layers)  $\leq F_{000}$  (all layers) because the previous flux estimates rely primarily on  $\text{CO}_2$  concentrations at deeper depths, and could miss high concentrations of  $\text{CO}_2$  produced in shallower layers.

When modeling soil respiration, typically a non-linear response function that also considers soil type is used (Bouma & Bryla, 2000; Yan et al., 2016, 2018). For the `neonSoilFlux` package, soil type is connected to the measurement of bulk density, which was characterized at each NEON site. This bulk density estimate is based on replicate samples collected from the site

354 megapit at a subset of soil horizons, with an estimated uncertainty of  $\pm 5\%$  (NEON, 2024c).  
355 Coarse fragment estimates also have very large uncertainties, but because the volume fraction  
356 tends to be low in surface soils it is unlikely to contribute much additional flux uncertainty.

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

370 We recognize that this gap-filling approach may lead to gap-filled values that are quite different  
371 from the actual values, such as an underestimate of soil moisture following rain events. Further  
372 extensions of the gap filling method could use more sophisticated gap-filling routines, similar to  
373 what is used for net ecosystem carbon exchange (Falge et al., 2001; Liu et al., 2023; Mariethoz et  
374 al., 2015; Moffat et al., 2007; Zhang et al., 2023). Additionally, since the deepest temperature  
375 and soil moisture sensors are located below the deepest CO<sub>2</sub> sensors at NEON sites, it is  
376 possible that excluding these deeper layers from consideration prior to analysis would lead to  
377 a reduced need for gap filling. Future iterations of the `neonSoilFlux` package may incorporate  
378 this as an option. The current gap-filling routine provides a consistent approach that can be

379 applied to each data stream, but further work may explore alternative gap-filling approaches.

## 380 6.2 Evaluation of flux-gradient approach at each site

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

388 The six locations studied provide a range of case studies that suggest different considerations  
389 may apply to different sites when applying the flux-gradient method. For example, the Santa  
390 Rita Experimental Range (SRER) is a desert site characterized by sandy soil, which also was  
391 the location of the highest field soil temperatures that we observed (Table 2). At SRER the  
392 flux across the top two layers ( $F_{110}$ ) produced a pattern of soil flux most consistent with the  
393 observed field data. The remaining methods  $F_{101}$ ,  $F_{011}$ , or  $F_{000}$  are derived from information  
394 taken from the deepest layer, which seems to have been decoupled from the surface layers both  
395 in terms of temperature and CO<sub>2</sub> concentration. This may be a general circumstance where  
396 there are large diurnal temperature extremes that rapidly change during the course of a day  
397 and overnight, leading to lags in the timing of when temperature increases propagate down to  
398 deeper soil layers.

399 Immediately prior to our visit to Konza Prairie (KONZ), that site that experienced a significant  
400 rain event that led to wet soils that gradually dried out over the course of our time there.  
401 This pulse of precipitation increased the soil CO<sub>2</sub> concentration at the top layer above the

402 concentrations in lower layers, leading to negative estimated flux values at the start of the  
403 field sampling period. In this case it was only when the soil began to return to a baseline level  
404 that the assumptions of the flux-gradient method were again met.

405 Both of the previous cases also provide context for the variable statistical comparisons between  
406 field-measured soil fluxes and `neonSoilFlux` outputs (Figure 6). When considering systematic  
407 deployment of this method across a measurement network, there are a number of independent  
408 challenges that require careful consideration. There are clear tradeoffs between (1) accuracy  
409 of modeled fluxes (defined here as closeness to field-measured  $F_S$  and the uncertainty reduc-  
410 tion factor  $\epsilon$ ), (2) precision (which could be defined by the signal to noise ratio), and (3) the  
411 choice of the diffusivity model (Section 4.2.2) or flux computation method (Section 4.2.3). A  
412 sensitivity analysis (Figure S2, Supplemental Information) found that flux output uncertainty  
413 was dominated by measurement uncertainty ( $T_S$ ,  $P$ ,  $SWC$ , or  $CO_2$ ) rather than by the dif-  
414 fusivity method used to compute soil flux. Notably, the  $F_{110}$  method was least sensitive to  
415 measurement uncertainty likely because it best aligns with the surface chamber measurement  
416 assumptions.

417 Finally, comparing the effects of different diffusivity estimation methods on the match between  
418 modeled and measured fluxes (Figure 5) highlights the sensitivity of  $F_{ijk}$  to diffusivity. The  
419 comparison between diffusivity estimates compared to field estimated diffusivity (Figure 7)  
420 demonstrates that site parameters can dictate which measure of diffusivity is most likely to  
421 be accurate in a given environmental context. Site-specific differences are largely a reflec-  
422 tion of differences in soil moisture across the sites (Table 1), as not all diffusivity estimation  
423 methods incorporate soil moisture equivalently. While we here have compares two approaches  
424 to calculate diffusivity (the Millington-Quirk and Marshall models), it may be valuable to  
425 evaluate other diffusivity models (e.g. the Moldrup model; Moldrup et al. (1999)) as well. Ul-  
426 timately the choice of a particular diffusivity model could be determined based on knowledge

427 of site-specific evaluations or a set of these models could be used to generate a model ensemble  
428 average as a means to trade precision for a more general approach.

429 **6.3 Recommendations for future method development**

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

440 These challenges notwithstanding, the method used here and made available in the  
441 `neonSoilFlux` R package has the potential to produce nearly continuous estimates of flux  
442 across all terrestrial NEON sites. These estimates are a significant improvement on available  
443 approaches to constrain the portion of ecosystem respiration attributable to the soil. This,  
444 in turn, also aids in our ability to understand the soil contribution to the net ecosystem flux  
445 measured at these sites using the co-located eddy flux towers.

<sup>446</sup> **7 Conclusions**

<sup>447</sup> We used the R package `neonSoilFlux` to estimate soil CO<sub>2</sub> fluxes with the flux-gradient  
<sup>448</sup> method using data from buried soil sensors at NEON terrestrial sites. We compared the  
<sup>449</sup> predicted fluxes to those measured directly using a field-based closed chamber approach. Soil  
<sup>450</sup> fluxes from `neonSoilFlux` were broadly effective at producing estimates of flux comparable  
<sup>451</sup> to those measured in the field using a chamber-based technique. However `neonSoilFlux`  
<sup>452</sup> outputs are quite sensitive to a number of issues, including: missing data (and thus gap-  
<sup>453</sup> filling of input measurement datasets), the selection of soil depths used to best calculate the  
<sup>454</sup> gradient (which may vary between sites), and finally the choice of method used for estimating  
<sup>455</sup> soil diffusivity. The flexibility of the `neonSoilFlux` package allows the user to evaluate each  
<sup>456</sup> of these issues with site-specific knowledge and contexts. Future refinements and subsequent  
<sup>457</sup> validation of `neonSoilFlux` outputs will feed forward into evaluating soil carbon fluxes broader  
<sup>458</sup> spatial scales to enhance understanding of the ways in which soils across diverse ecosystems  
<sup>459</sup> are responding to a changing climate.

<sup>460</sup> **Sources Cited**

- <sup>461</sup> Ayres, E., Reichle, R. H., Colliander, A., Cosh, M. H., & Smith, L. (2024). Validation of  
<sup>462</sup> Remotely Sensed and Modeled Soil Moisture at Forested and Unforested NEON Sites.  
<sup>463</sup> *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17,  
<sup>464</sup> 14248–14264. <https://doi.org/10.1109/JSTARS.2024.3430928>
- <sup>465</sup> Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the  
<sup>466</sup> atmosphere - the state and future of the eddy covariance method. *Global Change Biology*,  
<sup>467</sup> 20(12), 3600–3609. <https://doi.org/10.1111/gcb.12649>

- 468 Berenbaum, M. R., Carpenter, S. R., Hampton, S. E., Running, S. W., & Stanzione, D. C.  
469 (2015). *Report from the NSF BIO Advisory Committee Subcommittee on NEON Scope*  
470 *Impacts*.
- 471 Bond-Lamberty, B. (2018). New Techniques and Data for Understanding the Global Soil Res-  
472piration Flux. *Earth's Future*, 6(9), 1176–1180. <https://doi.org/10.1029/2018EF000866>
- 473 Bond-Lamberty, B., Ballantyne, A., Berryman, E., Fluet-Chouinard, E., Jian, J., Morris, K.  
474 A., Rey, A., & Vargas, R. (2024). Twenty Years of Progress, Challenges, and Opportuni-  
475ties in Measuring and Understanding Soil Respiration. *Journal of Geophysical Research: Bio-geosciences*, 129(2), e2023JG007637. <https://doi.org/10.1029/2023JG007637>
- 476 Bond-Lamberty, B., Christianson, D. S., Malhotra, A., Pennington, S. C., Sihi, D., AghaK-  
477ouchak, A., Anjileli, H., Altaf Arain, M., Armesto, J. J., Ashraf, S., Ataka, M., Baldocchi,  
478 D., Andrew Black, T., Buchmann, N., Carbone, M. S., Chang, S.-C., Crill, P., Curtis, P.  
479 S., Davidson, E. A., ... Zou, J. (2020). COSORE: A community database for continuous  
480 soil respiration and other soil-atmosphere greenhouse gas flux data. *Global Change Biology*,  
481 26(12), 7268–7283. <https://doi.org/10.1111/gcb.15353>
- 482 Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Bio-  
483geosciences*, 7(6), 1915–1926. <https://doi.org/10.5194/bg-7-1915-2010>
- 484 Bond-Lamberty, B., Wang, C., & Gower, S. T. (2004). A global relationship between the  
485 heterotrophic and autotrophic components of soil respiration? *Global Change Biology*,  
486 10(10), 1756–1766. <https://doi.org/10.1111/j.1365-2486.2004.00816.x>
- 487 Bouma, T. J., & Bryla, D. R. (2000). On the assessment of root and soil respiration for soils  
488 of different textures: Interactions with soil moisture contents and soil CO<sub>2</sub> concentrations.  
489 *Plant and Soil*, 227(1), 215–221. <https://doi.org/10.1023/A:1026502414977>
- 490 Chen, H., & Tian, H.-Q. (2005). Does a General Temperature-Dependent Q10 Model of Soil  
491 Respiration Exist at Biome and Global Scale? *Journal of Integrative Plant Biology*, 47(11),  
492 1288–1302. <https://doi.org/10.1111/j.1744-7909.2005.00211.x>

- 494 Davidson, E. A., Janssens, I. A., & Luo, Y. (2006). On the variability of respiration in  
495 terrestrial ecosystems: Moving beyond Q10. *Global Change Biology*, 12, 154–164. <https://doi.org/10.1111/j.1365-2486.2005.01065.x>
- 496
- 497 Desai, A. R., Murphy, B. A., Wiesner, S., Thom, J., Butterworth, B. J., Koupaei-Abyazani, N.,  
498 Muttaqin, A., Paleri, S., Talib, A., Turner, J., Mineau, J., Merrelli, A., Stoy, P., & Davis,  
499 K. (2022). Drivers of Decadal Carbon Fluxes Across Temperate Ecosystems. *Journal of*  
500 *Geophysical Research: Biogeosciences*, 127(12), e2022JG007014. <https://doi.org/10.1029/2022JG007014>
- 501
- 502 Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Chapman and  
503 Hall/CRC. <https://doi.org/10.1201/9780429246593>
- 504 Elshall, A. S., Ye, M., Pei, Y., Zhang, F., Niu, G.-Y., & Barron-Gafford, G. A. (2018). Relative  
505 model score: A scoring rule for evaluating ensemble simulations with application to micro-  
506 bial soil respiration modeling. *Stochastic Environmental Research and Risk Assessment*,  
507 32(10), 2809–2819. <https://doi.org/10.1007/s00477-018-1592-3>
- 508 Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G.,  
509 Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger,  
510 D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., ... Wofsy, S. (2001).  
511 Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural*  
512 *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)
- 513 Farrance, I., & Frenkel, R. (2012). *Uncertainty of Measurement: A Review of the Rules*  
514 *for Calculating Uncertainty Components through Functional Relationships*. *The Clinical*  
515 *Biochemist Reviews*, 33(2), 49–75.
- 516 Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer,  
517 P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz,  
518 J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., ...  
519 Zeng, J. (2025). Global Carbon Budget 2024. *Earth System Science Data*, 17(3), 965–1039.

- 520        <https://doi.org/10.5194/essd-17-965-2025>
- 521        Hamdi, S., Moyano, F., Sall, S., Bernoux, M., & Chevallier, T. (2013). Synthesis analysis  
522        of the temperature sensitivity of soil respiration from laboratory studies in relation to  
523        incubation methods and soil conditions. *Soil Biology and Biochemistry*, 58, 115–126. <https://doi.org/10.1016/j.soilbio.2012.11.012>
- 524        Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017).  
525        The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls.  
526        *Annual Review of Ecology, Evolution and Systematics*, 48(Volume 48, 2017), 419–445.  
527        <https://doi.org/10.1146/annurev-ecolsys-112414-054234>
- 528        Jian, J., Bailey, V., Dorheim, K., Konings, A. G., Hao, D., Shiklomanov, A. N., Snyder, A.,  
529        Steele, M., Teramoto, M., Vargas, R., & Bond-Lamberty, B. (2022). Historically inconsis-  
530        tent productivity and respiration fluxes in the global terrestrial carbon cycle. *Nature  
531        Communications*, 13(1), 1733. <https://doi.org/10.1038/s41467-022-29391-5>
- 532        Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N.,  
533        Manzon, J., Marchesi, R., Paredes, D., & Bond-Lamberty, B. (2021). A restructured and  
534        updated global soil respiration database (SRDB-V5). *Earth System Science Data*, 13(2),  
535        255–267. <https://doi.org/10.5194/essd-13-255-2021>
- 536        Jiang, J., Feng, L., Hu, J., Liu, H., Zhu, C., Chen, B., & Chen, T. (2024). Global soil  
537        respiration predictions with associated uncertainties from different spatio-temporal data  
538        subsets. *Ecological Informatics*, 82, 102777. <https://doi.org/10.1016/j.ecoinf.2024.102777>
- 539        Jobbágy, E. G., & Jackson, R. B. (2000). The Vertical Distribution of Soil Organic Carbon  
540        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)
- 541        Jurasinski, G., Koebisch, F., Guenther, A., & Beetz, S. (2022). *Flux: Flux Rate Calculation  
542        from Dynamic Closed Chamber Measurements*.
- 543        Liu, K., Li, X., Wang, S., & Zhang, H. (2023). A robust gap-filling approach for European  
544        soil respiration. *Soil Biology and Biochemistry*, 175, 109382. <https://doi.org/10.1016/j.soilbio.2023.109382>

- 546 Space Agency Climate Change Initiative (ESA CCI) soil moisture integrating satellite  
547 observations, model-driven knowledge, and spatiotemporal machine learning. *Hydrology*  
548 and *Earth System Sciences*, 27(2), 577–598. <https://doi.org/10.5194/hess-27-577-2023>
- 549 Lunch, C., Laney, C., Mietkiewicz, N., Sokol, E., Cawley, K., & Network), N. (National. E. O.  
550 (2025). *neonUtilities: Utilities for Working with NEON Data*.
- 551 Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011).  
552 Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*,  
553 21(5), 1429–1442. <https://doi.org/10.1890/09-1275.1>
- 554 Maier, M., & Schack-Kirchner, H. (2014). Using the gradient method to determine soil gas  
555 flux: A review. *Agricultural and Forest Meteorology*, 192–193, 78–95. <https://doi.org/10.1016/j.agrformet.2014.03.006>
- 557 Mariethoz, G., Linde, N., Jougnot, D., & Rezaee, H. (2015). Feature-preserving interpolation  
558 and filtering of environmental time series. *Environmental Modelling & Software*, 72, 71–76.  
559 <https://doi.org/10.1016/j.envsoft.2015.07.001>
- 560 Marshall, T. J. (1959). The Diffusion of Gases Through Porous Media. *Journal of Soil Science*,  
561 10(1), 79–82. <https://doi.org/10.1111/j.1365-2389.1959.tb00667.x>
- 562 Millington, R. J., & Shearer, R. C. (1971). Diffusion in aggregated porous media. *Soil Science*,  
563 111(6), 372–378.
- 564 Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G.,  
565 Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann,  
566 M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., & Stauch, V. J. (2007). Comprehensive  
567 comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and*  
568 *Forest Meteorology*, 147(3), 209–232. <https://doi.org/10.1016/j.agrformet.2007.08.011>
- 569 Moldrup, P., Olesen, T., Yamaguchi, T., Schjønning, P., & Rolston, D. E. (1999). Modeling  
570 diffusion and reaction in soils: 9. The Buckingham-Burdine-Campbell equation for gas  
571 diffusivity in undisturbed soil. *Soil Science*, 164(2), 75.

- 572 NEON. (2024a). *Barometric pressure (DP1.00004.001)*. National Ecological Observatory  
573 Network (NEON). <https://doi.org/10.48443/RT4V-KZ04>
- 574 NEON. (2024b). *Soil CO<sub>2</sub> concentration (DP1.00095.001)*. National Ecological Observatory  
575 Network (NEON). <https://doi.org/10.48443/E7GR-6G94>
- 576 NEON. (2024c). *Soil physical and chemical properties, Megapit (DP1.00096.001)*. National  
577 Ecological Observatory Network (NEON). <https://doi.org/10.48443/S6ND-Q840>
- 578 NEON. (2024d). *Soil temperature (DP1.00041.001)*. National Ecological Observatory Network  
579 (NEON). <https://doi.org/10.48443/Q24X-PW21>
- 580 NEON. (2024e). *Soil water content and water salinity (DP1.00094.001)*. National Ecological  
581 Observatory Network (NEON). <https://doi.org/10.48443/A8VY-Y813>
- 582 Norman, J. M., Kucharik, C. J., Gower, S. T., Baldocchi, D. D., Crill, P. M., Rayment, M.,  
583 Savage, K., & Striegl, R. G. (1997). A comparison of six methods for measuring soil-  
584 surface carbon dioxide fluxes. *Journal of Geophysical Research: Atmospheres*, 102(D24),  
585 28771–28777. <https://doi.org/10.1029/97JD01440>
- 586 Pedersen, A. R. (2024). *HMR: Flux Estimation with Static Chamber Data*.
- 587 Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown,  
588 K., & Vargas, R. (2017). The value of soil respiration measurements for interpreting and  
589 modeling terrestrial carbon cycling. *Plant and Soil*, 413(1), 1–25. <https://doi.org/10.1007/s11104-016-3084-x>
- 590
- 591 Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). *Using Bayesian Model  
592 Averaging to Calibrate Forecast Ensembles*. <https://doi.org/10.1175/MWR2906.1>
- 593 Rheault, K., Christiansen, J. R., & Larsen, K. S. (2024). goFlux: A user-friendly way to  
594 calculate GHG fluxes yourself, regardless of user experience. *Journal of Open Source  
595 Software*, 9(96), 6393. <https://doi.org/10.21105/joss.06393>
- 596 Sallam, A., Jury, W. A., & Letey, J. (1984). Measurement of Gas Diffusion Coefficient under  
597 Relatively Low Air-filled Porosity. *Soil Science Society of America Journal*, 48(1), 3–6.

- 598        <https://doi.org/10.2136/sssaj1984.03615995004800010001x>
- 599     Shao, J., Zhou, X., Luo, Y., Li, B., Aurela, M., Billesbach, D., Blanken, P. D., Bracho, R.,  
600           Chen, J., Fischer, M., Fu, Y., Gu, L., Han, S., He, Y., Kolb, T., Li, Y., Nagy, Z., Niu, S.,  
601           Oechel, W. C., ... Zhang, J. (2015). Biotic and climatic controls on interannual variability  
602           in carbon fluxes across terrestrial ecosystems. *Agricultural and Forest Meteorology*, 205,  
603           11–22. <https://doi.org/10.1016/j.agrformet.2015.02.007>
- 604     Shao, P., Zeng, X., Moore, D. J. P., & Zeng, X. (2013). Soil microbial respiration from  
605           observations and Earth System Models. *Environmental Research Letters*, 8(3), 034034.  
606           <https://doi.org/10.1088/1748-9326/8/3/034034>
- 607     Sihi, D., Gerber, S., Inglett, P. W., & Inglett, K. S. (2016). Comparing models of microbial–  
608           substrate interactions and their response to warming. *Biogeosciences*, 13(6), 1733–1752.  
609           <https://doi.org/10.5194/bg-13-1733-2016>
- 610     Tang, J., Baldocchi, D. D., Qi, Y., & Xu, L. (2003). Assessing soil CO<sub>2</sub> efflux using continuous  
611           measurements of CO<sub>2</sub> profiles in soils with small solid-state sensors. *Agricultural and Forest  
612           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)
- 613     Tang, J., Misson, L., Gershenson, A., Cheng, W., & Goldstein, A. H. (2005). Continuous  
614           measurements of soil respiration with and without roots in a ponderosa pine plantation  
615           in the Sierra Nevada Mountains. *Agricultural and Forest Meteorology*, 132(3), 212–227.  
616           <https://doi.org/10.1016/j.agrformet.2005.07.011>
- 617     Taylor, J. R. (2022). *An Introduction to Error Analysis: The Study of Uncertainties in Physical  
618           Measurements, Third Edition* (3rd ed.). University Science Press.
- 619     Wilson, S. J., Bond-Lamberty, B., Noyce, G., Bittencourt Peixoto, R., & Megonigal, J. P.  
620           (2024). Fluxfinder: An R Package for Reproducible Calculation and Initial Processing of  
621           Greenhouse Gas Fluxes From Static Chamber Measurements. *Journal of Geophysical Re-  
622           search: Biogeosciences*, 129(11), e2024JG008208. <https://doi.org/10.1029/2024JG008208>
- 623     Yan, Z., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S., Liu, C., & Liu, C. (2018).

- 624 A moisture function of soil heterotrophic respiration that incorporates microscale processes.
- 625 *Nature Communications*, 9(1), 2562. <https://doi.org/10.1038/s41467-018-04971-6>
- 626 Yan, Z., Liu, C., Todd-Brown, K. E., Liu, Y., Bond-Lamberty, B., & Bailey, V. L. (2016).
- 627 Pore-scale investigation on the response of heterotrophic respiration to moisture conditions
- 628 in heterogeneous soils. *Biogeochemistry*, 131(1), 121–134. <https://doi.org/10.1007/s10533-016-0270-0>
- 630 Zhang, R., Kim, S., Kim, H., Fang, B., Sharma, A., & Lakshmi, V. (2023). Temporal
- 631 Gap-Filling of 12-Hourly SMAP Soil Moisture Over the CONUS Using Water Balance
- 632 Budgeting. *Water Resources Research*, 59(12), e2023WR034457. <https://doi.org/10.1029/2023WR034457>
- 633 Zhao, J. (2019). FluxCalR: A R package for calculating CO<sub>2</sub> and CH<sub>4</sub> fluxes from static
- 634 chambers. *Journal of Open Source Software*, 4(43), 1751. <https://doi.org/10.21105/joss.01751>
- 635 Zobitz, J., Ayres, E., O'Rourke, K., Werbin, Z., Lee, L., Abdi, R., Mehmeti, D., & Xiong, L.
- 636 (2024). *neonSoilFlux: Compute Soil Carbon Fluxes for the National Ecological Observatory*
- 637 *Network Sites*.