

1 A direct comparison between field-measured

2 and sensor-based estimates of soil carbon

3 dioxide flux across six National Ecological

4 Observatory Network sites enabled by the

5 neonSoilFlux R package

6 John Zobitz¹ Ed Ayres² Zoey Werbin³ Ridwan Abdi¹

7 Natalie Ashburner-Wright⁴ Lillian Brown⁴

8 Ryan Frink-Sobierajski⁴ Lajntxiag Lee¹ Dijonë Mehmeti¹

9 Christina Tran⁴ Ly Xiong¹ Naupaka Zimmerman⁴

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

11 ² National Ecological Observatory Network, 1685 38th Street, Suite 100, Boulder, CO 80301

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

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

¹⁴ **Acknowledgments**

¹⁵ JZ acknowledges Kathleen O'Rourke for code development. NZ thanks technical staff at
¹⁶ USF for support with field gear assembly and shipping. We thank the NEON field staff
¹⁷ and assignable assets teams for facilitating each of the six NEON site visits. We are grateful
¹⁸ to LI-COR technical staff for helpful discussions about optimal sampling methods. This work
¹⁹ was supported by NSF DEB grant #2017829 awarded to JZ, and NSF DEB grant #2017860
²⁰ awarded to NZ.

²¹ **Conflict of Interest Statements**

²² None of the authors have a financial, personal, or professional conflict of interest related to
²³ this work.

²⁴ **Author Contributions**

²⁵ Conceptualization: JZ, NZ; Methodology: EA, JZ, NZ; Software: JZ, NZ, ZW, E A, DM, RA,
²⁶ LX, LL; Validation: JZ, NZ; Formal Analysis: JZ, NZ, DM, RA, LX, LL; Investigation: JZ,
²⁷ NZ, RF-S, CT, NA-W, LB; Resources: JZ, NZ; Data curation: JZ, NZ, DM, LX; Writing
²⁸ – original draft: JZ, NZ; Writing – review and editing: JZ, NZ, ZW, EA, CT, DM, LX,;
²⁹ Visualization: JZ, NZ, DM, RA, LX; Supervision: JZ; NZ; Project Administration: JZ; NZ;
³⁰ Funding Acquisition: JZ; NZ

³¹ **Data Availability**

³² Data available from the Zenodo LINK <http://dx.doi.org/10.5061/dryad.41qh7> (Kiere & Drummond 2016)."

³⁴ **1 Abstract**

³⁵ A key component of constraining the uncertainty of the terrestrial carbon sink is quantification
³⁶ of terrestrial soil carbon fluxes, which vary across time and ecosystem type. One method for
³⁷ the estimation of these fluxes and their associated uncertainties is the flux gradient method,
³⁸ which can be calculated via a variety of existing approaches. Robust estimation of soil carbon
³⁹ fluxes on a sub-daily level requires measurements of soil CO₂ concentration, water content,
⁴⁰ temperature, and other environmental measurements and soil properties. These data are
⁴¹ publicly available from the National Ecological Observatory Network at sites spanning a range
⁴² of 20 different ecoclimatic domains across the continental United States, Puerto Rico, Alaska,
⁴³ and Hawai'i. We present an R software package (`neonSoilFlux`) that acquires NEON soil
⁴⁴ environmental data and computes soil carbon flux at a half-hourly time step at a user-specified
⁴⁵ NEON site and month in a tidy data format. To validate the computed fluxes, we visited six
⁴⁶ focal NEON sites and measured soil carbon fluxes using a closed-dynamic chamber approach.
⁴⁷ The validation confirmed that a primary challenge in reducing soil carbon flux uncertainty is
⁴⁸ correctly characterizing diffusivity and soil water content across the soil profile. Outputs from
⁴⁹ the `neonSoilFlux` package contribute to existing databases of soil carbon flux measurements,
⁵⁰ providing near real-time estimates of a critical component of the terrestrial carbon cycle.

⁵¹ **1.1 Keywords**

⁵² Soil carbon, carbon dioxide, flux gradient, carbon cycle, field validation, soil respiration, ecosys-
⁵³ tem variability, diffusion

⁵⁴ **2 Data for peer review**

⁵⁵ Anonymous data and code for peer review is available here: [LINK](#)

⁵⁶ **3 Introduction**

⁵⁷ Soils contain the largest reservoir of terrestrial carbon (Jobbágy & Jackson, 2000). A critical
⁵⁸ component of this reservoir is soil organic matter, the accumulation of which is influenced
⁵⁹ by biotic factors such as above-ground plant inputs (Jackson et al., 2017). These inputs in
⁶⁰ turn are influenced by environmental factors such as growing season length, temperature, and
⁶¹ moisture (Desai et al., 2022), which also affect the breakdown of soil organic matter and its
⁶² return to the atmosphere. Across heterogeneous terrestrial landscapes, the interplay between
⁶³ these biotic and abiotic factors influence the size of the soil contribution to the terrestrial
⁶⁴ carbon sink (Friedlingstein et al., 2023). However, the heterogeneity of these processes across
⁶⁵ diverse ecosystems in the context of rapid environmental change leads to large uncertainty in
⁶⁶ the magnitude of this sink in the future, and thus a pressing need to quantify changes in soil
⁶⁷ carbon pools and fluxes across scales.

⁶⁸ Ecological observation networks such as the United States' National Ecological Observatory
⁶⁹ Network (NEON) and others (e.g. FLUXNET or the Integrated Carbon Observation System)
⁷⁰ present a significant advancement in the nearly continuous observation of biogeochemical pro-
⁷¹ cesses at the continental scale. Notably, at 47 terrestrial sites across the continental United
⁷² States, NEON provides half-hourly measurements of soil CO₂ concentration, temperature,
⁷³ and moisture at different vertical depths. Each of these NEON sites also encompasses mea-
⁷⁴ surements of the cumulative sum of all ecosystem carbon fluxes in an airshed using the eddy
⁷⁵ covariance technique (Balderuppi, 2014). Soil observations provided by NEON are on the same

76 timescale and standardized with eddy covariance measurements from FLUXNET. These types
77 of nearly continuous observational data (NEON and FLUXNET) can be used to reconcile dif-
78 ferences between model-derived or data-estimated components of ecosystem carbon flux (Jian
79 et al., 2022; Luo et al., 2011; Phillips et al., 2017; J. Shao et al., 2015; P. Shao et al., 2013;
80 Sihi et al., 2016).

81 Estimated or observed soil carbon fluxes are a key metric for understanding change in soil
82 carbon pools over time (Bond-Lamberty et al., 2024). A soil carbon flux to the atmosphere
83 (F_S , units $\mu\text{mol m}^{-2} \text{s}^{-1}$), represents the aggregate process of transfer of soil CO_2 to the
84 atmosphere from physical and biological processes (e.g. diffusion and respiration). Soil carbon
85 fluxes can be assumed to encompass soil carbon respiration from autotrophic or heterotrophic
86 sources (Davidson et al., 2006), typically assumed to be static across the soil biome and
87 modeled with a exponential Q_{10} paradigm (Bond-Lamberty et al., 2004; Chen & Tian, 2005;
88 Hamdi et al., 2013).

89 One method by which F_S is measured in the field is through the use of soil chambers in a closed,
90 well-mixed system (Norman et al., 1997) with headspace trace gas concentrations measured
91 with an infrared gas analyzer (IRGA). F_S can also be estimated from soil CO_2 measurements
92 at different depths in the soil using the flux-gradient method (Maier & Schack-Kirchner, 2014).
93 This method is an approach that uses conservation of mass to calculate flux at a vertical soil
94 depth z at steady state by applying Fick's law of diffusion. A simplifying assumption for the
95 flux-gradient method is that there is no mass transfer in the other spatial dimensions x and y
96 (Maier & Schack-Kirchner, 2014). The diffusivity profile, a key component of this calculation,
97 varies across the soil depth as a function of soil temperature, soil volumetric water content,
98 atmospheric air pressure, and soil bulk density (Millington & Shearer, 1971; Moldrup et al.,
99 1999; Sallam et al., 1984).

100 Databases such as the Soil Respiration Database (SRDB) or the Continuous Soil Respiration

101 Database (COSORE) add to the growing network of resources for making collected observa-
102 tions of soil fluxes available to other workers (Bond-Lamberty, 2018; Bond-Lamberty et al.,
103 2020; Bond-Lamberty & Thomson, 2010; Jian et al., 2021; Jiang et al., 2024). However, these
104 databases currently encompass primarily direct soil measurements of fluxes (i.e. those using
105 methods like the closed-chamber method described above). Currently, NEON provides all
106 measurements to calculate F_S from Fick's law, but soil flux as a derived data product was
107 descoped from the initial network launch due to budget constraints (Berenbaum et al., 2015).
108 Deriving estimates of F_S using continuous sensor data across NEON sites thus represents a
109 high priority.

110 This study describes an R software package, `neonSoilFlux`, that can be used to derive a
111 standardized estimate of F_S at all terrestrial NEON sites. After calculating these flux estimates,
112 we then validated them against direct chamber-based field observations of soil carbon dioxide
113 flux from a subset of terrestrial NEON sites spanning six states.

114 Key objectives of this study are to:

- 115 1. Apply the flux-gradient method to estimate soil CO₂ flux from continuous sensor mea-
116 surements across NEON sites.
- 117 2. Benchmark estimated soil carbon fluxes against field measurements (e.g. direct chamber
118 measurements of soil flux).
- 119 3. Identify sources of error in the flux-gradient approach across diverse sites in order to
120 guide future work.

₁₂₁ **4 Materials and Methods**

₁₂₂ **4.1 Field methods**

₁₂₃ **4.1.1 Focal NEON Sites**

₁₂₄ In order to acquire field data to validate model predictions of flux, we selected six terrestrial
₁₂₅ NEON sites for analysis. We conducted field measurement campaigns at these sites, which
₁₂₆ span a range of environmental gradients and terrestrial domains (Table 1). SJER, SRER, and
₁₂₇ WREF were visited during May and June of 2022, and WOOD, KONZ, and UNDE during
₁₂₈ May and June of 2024.

₁₂₉ Over the course of two field campaigns in 2022 and 2024, we conducted week-long visits at
₁₃₀ each site. In consultation with NEON field staff, we first selected a specific plot in the soil
₁₃₁ sampling array to maximize the concurrent availability of sensor data.

₁₃₂ **4.1.2 Soil collar placement**

₁₃₃ Either one (2022 sampling campaign) or two (2024 sampling campaign) PVC soil collars (20.1
₁₃₄ cm inside diameter) were installed in close proximity to the permanent NEON soil sensors at
₁₃₅ each site (Figure 1). The soil plot where measurements were taken was chosen at each site
₁₃₆ in consultation with NEON staff to maximize likelihood of quality soil sensor measurements
₁₃₇ during the duration of the IRGA measurements at each site. After installation, collar(s) were
₁₃₈ left to equilibrate for approximately 24 hours prior to measurements being taken.

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

¹⁴⁰ In 2022, we then made measurements of flux on an hourly interval for 8 hours each day.
¹⁴¹ Measurements were taken from roughly 8 am to 4 pm, with the time interval selected to
¹⁴² capture the majority of the diurnal gradient of soil temperature each day. These measurements
¹⁴³ were made using a LI-6800 infrared gas analyzer instrument (LI-COR Environmental, Lincoln,
¹⁴⁴ NE) fitted with a soil chamber attachment (attachment 6800-09). In 2024, we again used
¹⁴⁵ the same LI-6800 instrument, but made half-hourly measurements over an approximately 8
¹⁴⁶ hour period. In addition, we also installed a second collar and used a second instrument, an
¹⁴⁷ LI-870 CO₂ IRGA, connected to an automated robotic chamber (LI-COR chamber 8200-104)
¹⁴⁸ controlled by an LI-8250 multiplexer, to make automated measurements. The multiplexer was
¹⁴⁹ configured to take half-hourly measurements 24 hours a day for the duration of our sampling
¹⁵⁰ bout at each site. Each instrument was paired with a soil temperature and moisture probe
¹⁵¹ (Stevens HydraProbe, Stevens Water, Portland, OR) that was used to make soil temperature
¹⁵² and moisture measurements concurrent with the CO₂ flux measurements. Chamber volumes
¹⁵³ were set by measuring collar offsets at each site. System checks were conducted daily for the
¹⁵⁴ LI-6800 and weekly for the LI-8250. Instruments were factory calibrated before each field
¹⁵⁵ season.

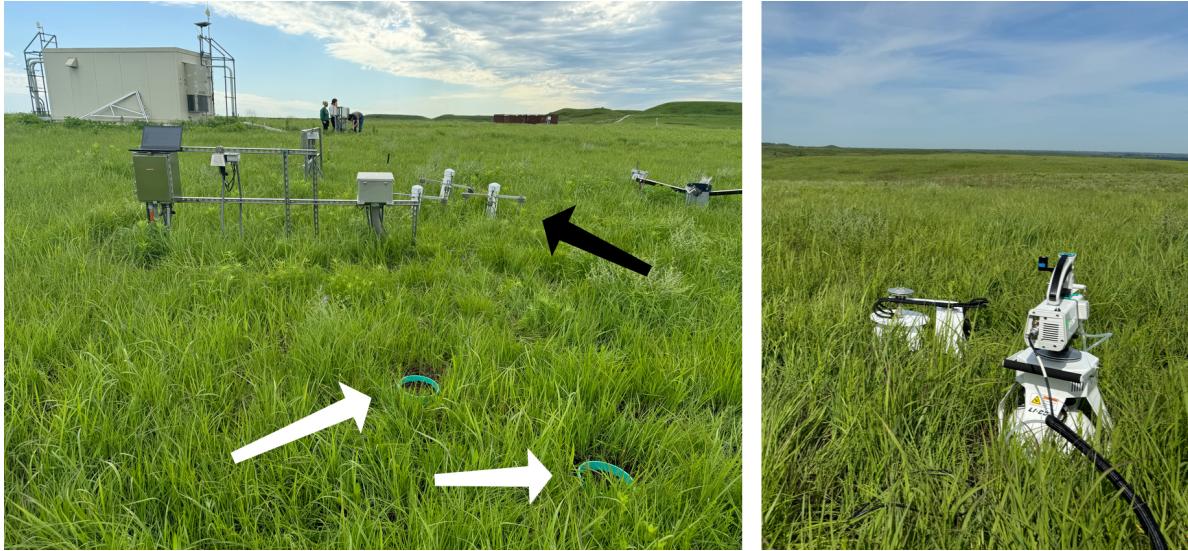


Figure 1: Spatial layout of field sampling using a closed-dynamic chamber setup at a representative NEON site (KONZ). Left image shows collars (white arrows) and permanent soil sensor installation (black arrow) and right image shows the LI-6800 (foreground) and LI-8200-104 (background) instruments placed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

Table 1: Listing of NEON sites studied for field work and analysis. $\overline{T_S}$: average soil temperature during field measurements. \overline{SWC} : average soil water content during field measurements. Soil plot refers to the particular location in the soil sensor array (denoted as HOR by NEON) where field measurements were made.

Site (NEON site ID)	Location	Ecosystem type	Mean annual tempera- ture	$\overline{T_S}$ (°)	Mean annual precipita- tion	\overline{SWC} (%)	Field measure- ment dates	Soil plot
Santa	31.91068,	Shrubland	19.3°C	47.6°	346 mm	4.0%	29 May	004
Rita	-						2024 - 01	
Experi- mental Range (SRER)	110.83549						June 2024	

Table 1: Listing of NEON sites studied for field work and analysis. \bar{T}_S : average soil temperature during field measurements. \bar{SWC} : average soil water content during field measurements. Soil plot refers to the particular location in the soil sensor array (denoted as HOR by NEON) where field measurements were made.

Site (NEON site ID)	Location	Ecosystem type	Mean annual tempera- ture	\bar{T}_S (°)	Mean annual precipita- tion	\bar{SWC} (%)	Field measure- ment dates	Soil plot
San Joaquin Experimental Range (SJER)	37.10878, -	Oak woodland	16.4°C	41.7°	540 mm	1.2%	01 June 2022 - 04	005
Wind River Experimental Forest (WREF)	45.82049, -	Evergreen forest	9.2°C	15.3°	2225 mm	27.2%	07 June 2022	001
Chase Lake National Wildlife Refuge (WOOD)	121.95191	Restored prairie	4.9°C	14.9°	495 mm	14.9%	03 June 2024 - 09	001
Konza Prairie Biological Station (KONZ)	47.1282, -	Tallgrass prairie	12.4°C	23.4°	870 mm	23.4%	29 May 2024 - 01	001
	99.241334	grassland					June 2024	
	96.563075						June 2024	

Table 1: Listing of NEON sites studied for field work and analysis. \bar{T}_S : average soil temperature during field measurements. \bar{SWC} : average soil water content during field measurements. Soil plot refers to the particular location in the soil sensor array (denoted as HOR by NEON) where field measurements were made.

Site (NEON site ID)	Location	Ecosystem type	Mean annual tempera- ture	\bar{T}_S (°)	Mean annual precipita- tion	\bar{SWC} (%)	Field measure- ment dates	Soil plot
University of Notre Dame Environmental Research Center (UNDE)	46.23391, - 89.537254	Deciduous forest	4.3°	13.0°	802 mm	13.0%	22 May 2024 - 25 May 2024	004

156 4.1.4 Post-collection processing of data

157 We used LI-COR SoilFluxPro software (v 5.3.1) to assess the data after collection and to inform
 158 sampling parameters. We checked appropriateness of dead band and measurement durations
 159 using built-in evaluation tools. Based on this, the deadband period was set for 30-40 seconds,
 160 depending on the site, and the measurement duration was 180 seconds with a 30 second pre-
 161 purge and a 30 second post-purge. We also assessed the R^2 of linear and exponential model
 162 fits to measured CO_2 to verify measurement quality.

163 4.2 neonSoilFlux R package

164 We developed an R package (`neonSoilFlux`; <https://CRAN.R-project.org/package=neonSoilFlux>)
 165 to compute half-hourly soil carbon fluxes and uncertainties from NEON data. The objective

166 of the `neonSoilFlux` package is a unified workflow (Figure 2) for soil data acquisition and
167 analysis that supplements the existing data acquisition R package `neonUtilities` (LINK TO
168 BE ADDED AFTER PEER REVIEW).

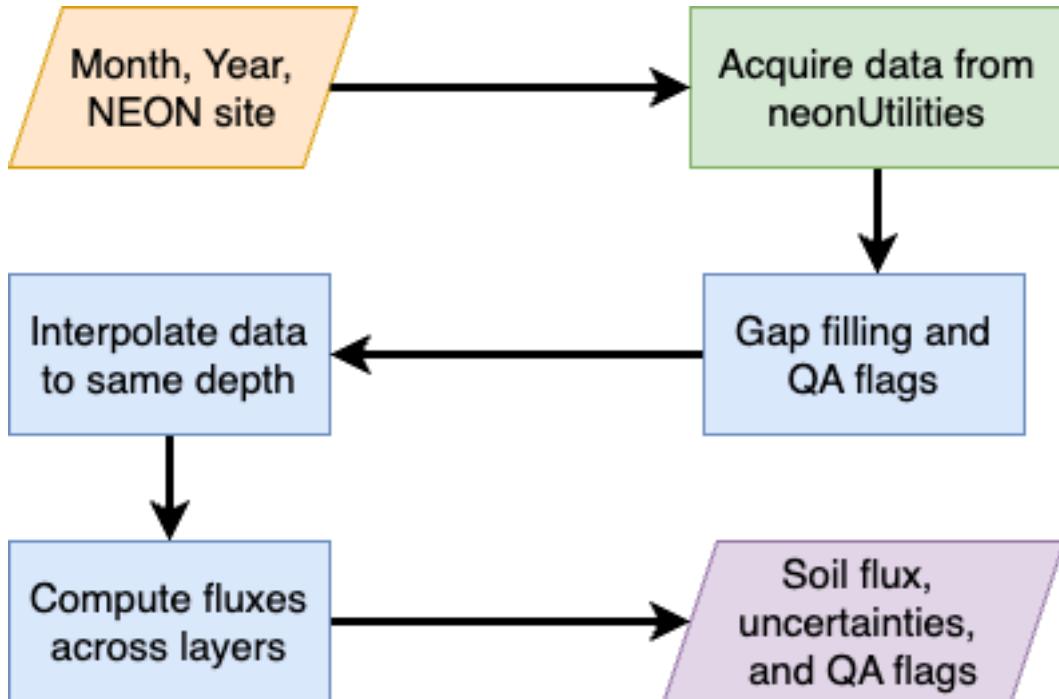


Figure 2: Diagram of `neonSoilFlux` R package. For a given month, year, and NEON site (orange parallelogram), the package acquires all relevant data to compute F_S using the `neonUtilities` R package (green rectangle). Data are gap-filled according to reported QA flags and interpolated to the same measurement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

169 At a given NEON observation there are five replicate soil plots, each with measurements of
170 soil CO₂ concentration, soil temperature, and soil moisture at different depths (Figure 3). The
171 `neonSoilFlux` package acquires measured soil water content (National Ecological Observatory
172 Network (NEON), 2024e), soil CO₂ concentration (National Ecological Observatory Network
173 (NEON), 2024b), barometric pressure from the nearby tower (National Ecological Observa-
174 tory Network (NEON), 2024a), soil temperature (National Ecological Observatory Network

175 (NEON), 2024d), and soil properties (e.g. bulk density) (National Ecological Observatory Net-
 176 work (NEON), 2024c). The static soil properties were collected from a nearby soil pit during
 177 site characterization and are assumed to be constant at each site.

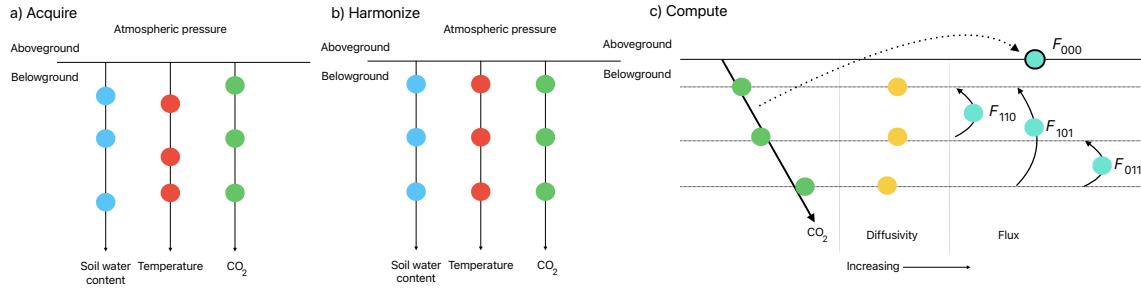


Figure 3: Model diagram for data workflow for the `neonSoilFlux` R package. a) Acquire: Data are obtained from given NEON location and horizontal sensor location, which includes soil water content, soil temperature, CO_2 concentration, and atmospheric pressure. All data are screened for quality assurance; if gap-filling of missing data occurs, it is flagged for the user. b) Any belowground data are then harmonized to the same depth as CO_2 concentrations using linear regression. c) The flux across a given depth is computed via Fick's law, denoted with F_{ijk} , where i , j , or k are either 0 or 1 denoting the layers the flux is computed across ($i = \text{closest to surface}$, $k = \text{deepest}$). F_{000} represents a flux estimate where the gradient dC/dz is the slope of a linear regression of CO_2 with depth.

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

183 The second step is harmonizing the data to compute soil fluxes across soil layers. This step
 184 consists of three different actions (blue rectangles in Figure 2 and Panel b in Figure 3). If a
 185 given observation by NEON is reported as not passing a quality assurance check, we applied
 186 a gap filling method to replace that measurement with its monthly mean at that same depth
 187 (Section 4.2.1). Belowground measurements of soil water and soil temperature are then inter-

188 polated to the same depth as soil CO₂ measurements. The diffusivity (Section 4.2.2) and soil
189 flux across different soil layers (Section 4.2.3) are then computed.

190 The third and final step is computing a surface soil flux through extrapolation to the sur-
191 face (purple parallelogram in Figure 2 and Panel c in Figure 3). Uncertainty on a soil flux
192 measurement is computed through quadrature. An aggregate quality assurance (QA) flag
193 for each environmental measurement is also reported, representing if any gap-filled measure-
194 ments were used in the computation of a soil flux. Within the soil flux-gradient method,
195 several different approaches can be used to derive a surface flux (Maier & Schack-Kirchner,
196 2014); the `neonSoilFlux` package reports four different possible values for soil surface flux
197 (Section 4.2.3).

198 4.2.1 Gap-filling routine

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

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

- 207 1. Randomly sample from the uncertainty and bias independently: σ_j and the bias \mathbf{b}_k (not
208 necessarily the same sample).
- 209 2. Generate a vector \mathbf{n} of length N , where \mathbf{n}_i is a random sample from a normal distribution
210 with mean m_i and standard deviation σ_j . Since $M < N$, values from \mathbf{m} will be reused.

211 3. With these N random samples, $\bar{y}_i = \bar{x} + \vec{b}_k$ and s_i is the sample standard deviation of \bar{x} .

212 We expect that $s_i \approx \vec{\sigma}_j$.

213 4. The reported monthly mean and standard deviation are then computed $\bar{\bar{y}}$ and \bar{s} . Mea-

214 surements and uncertainties that did not pass the QA check are then substituted with

215 $\bar{\bar{y}}$ and \bar{s} .

216 This gap-filling method described here provides a consistent approach for each data stream,

217 however we recognize that other gap-filling alternatives may be warranted for longer-term gaps

218 (e.g. such as correlations with other NEON measurement levels and soil plots), or measure-

219 ment specific gap-filling routines. We discuss the effect of gap-filling on our measurements in

220 Section 6.

221 **4.2.2 Soil diffusivity**

222 Soil diffusivity D_a at a given measurement depth is the product of the diffusivity in free air

223 $D_{a,0}$ ($\text{m}^2 \text{ s}^{-1}$) and the tortuosity ξ (no units) (Millington & Shearer, 1971).

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

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

225 where T_i is soil temperature ($^\circ\text{C}$) at depth i (National Ecological Observatory Network

226 (NEON), 2024d) and P surface barometric pressure (kPa) (National Ecological Observatory

227 Network (NEON), 2024a).

228 Previous studies by Sallam et al. (1984) and Tang et al. (2003) demonstrated the sensitivity

229 of modeled F_S depending on the tortuosity model used to compute diffusivity. At low soil

230 water content, the choice of tortusosity model may lead to order of magnitude differences in
 231 D_a , which in turn affect modeled F_S . The `neonSoilFlux` package uses two different models
 232 for ξ , representing the extremes reported in Sallam et al. (1984). The first approach uses the
 233 Millington-Quirk model for diffusivity, Equation 2 (Millington & Shearer, 1971):

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

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

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

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

243 The second approach to calculate ξ is the Marshall model (Marshall, 1959), where $\xi = \phi^{1.5}$,
 244 with ϕ defined from Equation 3.

245 **4.2.3 Soil flux computation**

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

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

248 where D_a is the diffusivity ($\text{m}^2 \text{s}^{-1}$) and $\frac{dC}{dz}$ is the gradient of CO_2 molar concentration
249 ($\mu\text{mol m}^{-3}$, so the gradient has units of $\mu\text{mol m}^{-3} \text{m}^{-1}$). The soil surface flux is theoretically
250 defined by applying Equation 4 to measurements collected at the soil surface and directly
251 below the surface. Measurements of soil temperature, soil water content, and soil CO_2 molar
252 concentration across the soil profile allow for application of Equation 4 across different soil
253 depths. Each site had three measurement layers, so we denote the flux between which two
254 layers as a three-digit subscript F_{ijk} with indicator variables i , j , and k indicate if a given
255 layer was used (written in order of increasing depth), according to the following:

- 256 • F_{000} is a surface flux estimate using the intercept of the linear regression of D_a with
257 depth and the slope from the linear regression of CO_2 with depth (which represents $\frac{dC}{dz}$
258 in Fick's Law). Tang et al. (2003) used this approach to compute fluxes in an oak-grass
259 savannah.
- 260 • F_{110} , F_{011} are fluxes across the two most shallow layers and two deepest layers respec-
261 tively. The diffusivity used in Fick's Law is always at the deeper measurement layer.
262 When used as a surface flux estimate we assume CO_2 remains constant above this flux
263 depth.
- 264 • F_{101} is a surface flux estimate using linear extrapolation using concentration measure-
265 ments between the shallowest and deepest measurement layer. Hirano et al. (2003) and

266 Tang et al. (2005) used an approach similar to F_{101} in a temperate deciduous broadleaf
267 forest and ponderosa pine forest respectively.

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

269 **4.3 Post processing evaluation**

270 Following collection of field measurements and calculation of the soil fluxes from `neonSoilFlux`
271 package, we compared measured F_S based on closed-dynamic chamber measurements with the
272 LI-COR instruments to a given soil flux calculation from `neonSoilFlux` for each site and flux
273 computation method. Statistics included the associated R^2 value, root mean squared error
274 (RMSE), and signal to noise ratio (SNR), defined as the ratio of a modeled soil flux (F_{ijk})
275 from `neonSoilFlux` to its quadrature uncertainty (σ_{ijk}).

276 We observed that the range of values (e.g. $F_{ijk} \pm \sigma_{ijk}$) was much larger than the measured
277 field flux. We evaluated $|F_S - F_{ijk}| < (1 - \epsilon)\sigma_{ijk}$, where F_S is a measured field soil flux from
278 the LI-COR 6800 (as the LI-COR 870/8250 was used at only three sites in 2024 but the 6800
279 was used at all sites in both years). The parameter ϵ was an uncertainty reduction factor to
280 evaluate how much the quadrature uncertainty could be reduced while maintaining precision
281 between modeled F_{ijk} and measured F_S .

282 Finally, for a half-hourly interval we also computed a *post hoc* D_a using the LI-COR flux along
283 with the CO₂ surface gradient reported by NEON using the measurement levels closest to the
284 surface.

285 **5 Results**

286 The timeseries of the measured fluxes from the LI-COR 6800 and 870/8250 were compared to
 287 modeled soil fluxes from the `neonSoilFlux` R package (Figure 4). We also assessed year-long
 288 estimated flux time series and compared those field measurements for each measurement site
 289 (Figure 5). Results are reported in local time. Positive values of the flux indicate that there
 290 is a flux moving towards the surface. For ease of clarity the fluxes at F_{111} and F_{000} are only
 291 shown in the top row (surface), followed by the fluxes at individual separate layers (F_{100} , F_{010} ,
 292 F_{001}). Overall, with the exception of WREF and SRER (discussed later) the computed fluxes
 293 determined using a variety of plausible methods spanned the measured field fluxes.

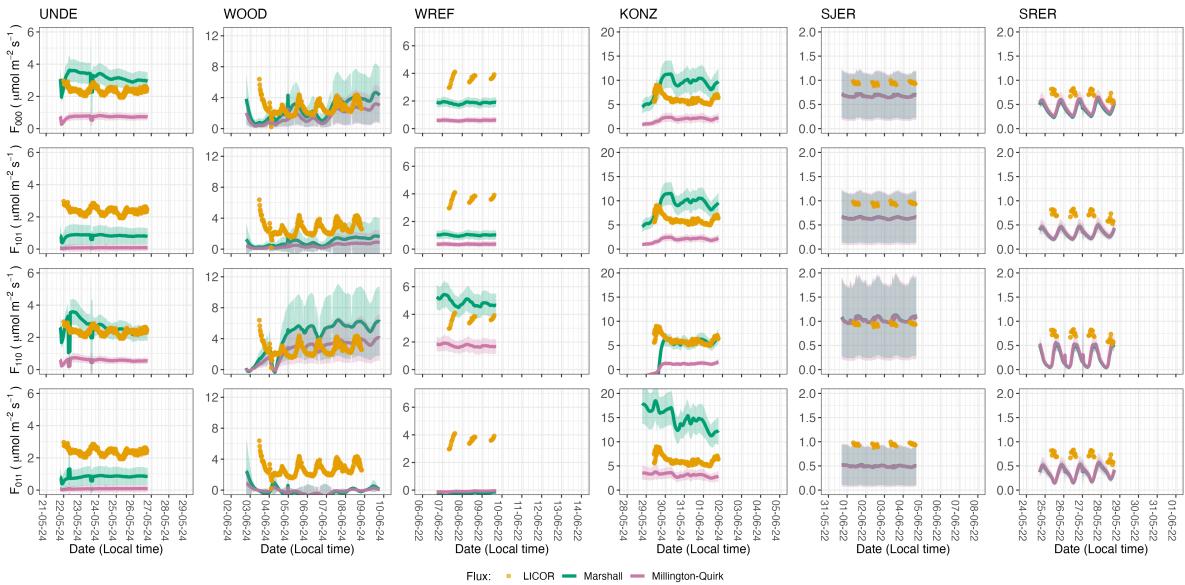


Figure 4: Timeseries of both measured F_S (yellow circles) and modeled soil fluxes (green or purple lines) by the `neonSoilFlux` R package. Fluxes from the `neonSoilFlux` R package are separated by the diffusivity model used (Millington-Quirk or Marshall, Section 4.2.2). Vertical axis labels in the first 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.

294 Table 2 compares statistics of the modeled soil fluxes to field F_S at each site. As `neonSoilFlux`

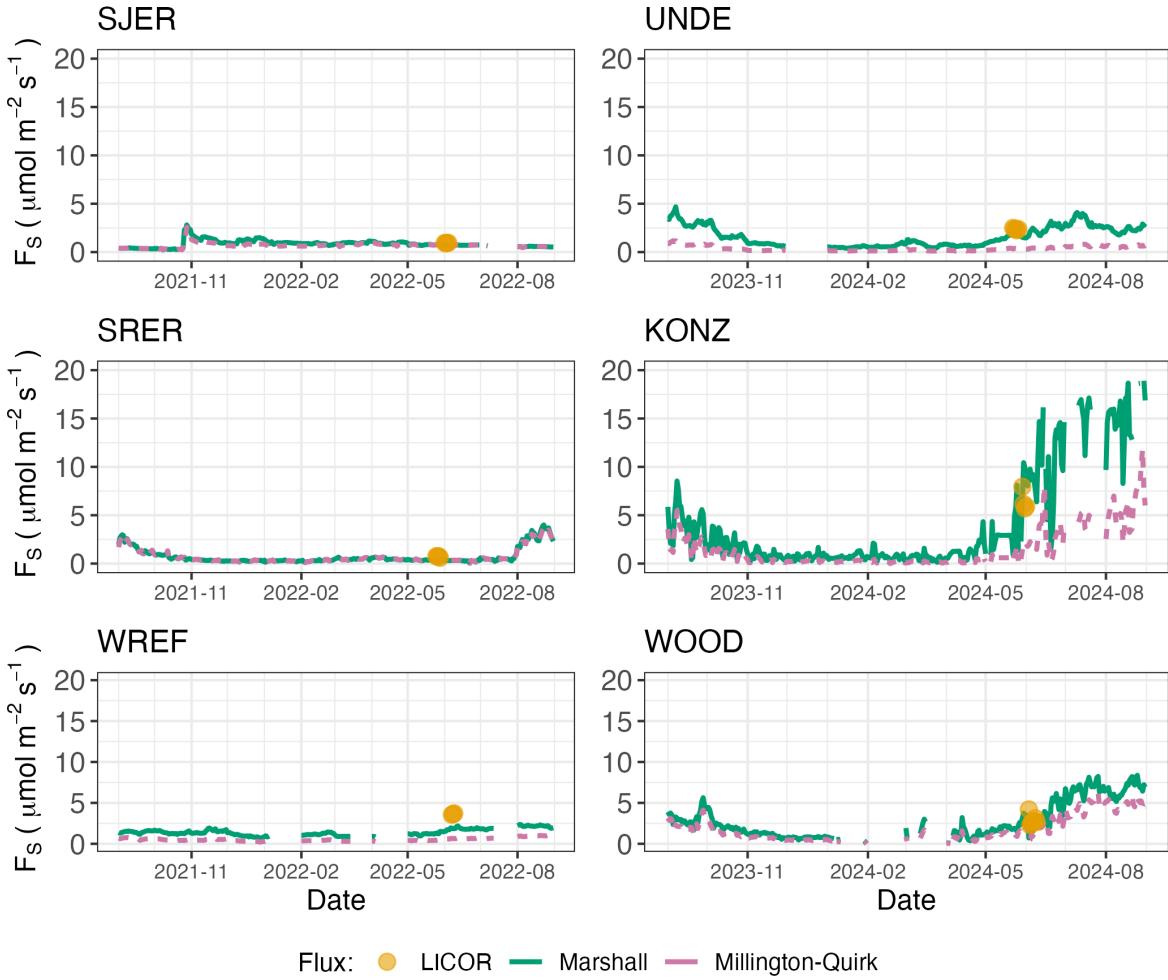
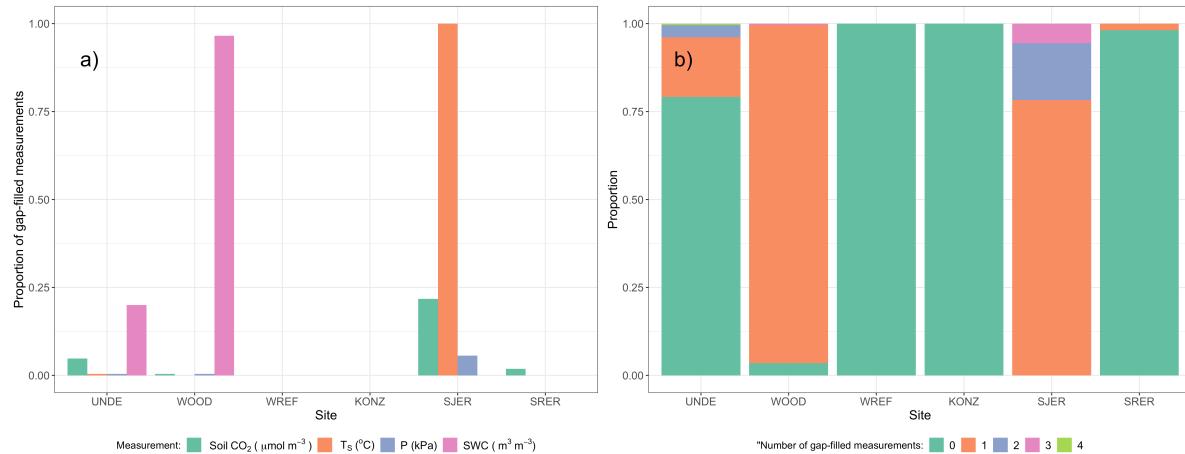


Figure 5: Timeseries of both daily-averaged field F_S (yellow circles) and daily ensemble averaged soil fluxes (green or purple lines) by the `neonSoilFlux` R package, separated by the diffusivity model used (Millington-Quirk or Marshall, Section 4.2.2). The time-series of modeled fluxes are a daily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

295 models half-hourly fluxes, we considered all measured fluxes over the duration of the half-hourly
 296 interval when computing the normalized root mean square error (NRMSE), R^2 , at each of the
 297 flux methods (F_{000} , F_{110} , F_{011} , F_{101}) for the diffusivity models studied (Millington-Quirk or
 298 Marshall). MORE HERE

299 For a given half-hourly time period, the `neonSoilFlux` package assigns a QA flag for a mea-
 300 surement if more one values across all measurement depths uses gap-filled data (Section 4.2.1).
 301 Panel a of Figure 6 reports the distribution for all input environmental measurements at each
 302 site when field measurements were made. Soil fluxes are computed from 4 different types of
 303 input measurements (T_S , SWC , P , and CO_2), any of which could have a QA flag in a half-
 304 hourly interval. Panel b of Figure 6 displays at each site the distribution of the number of
 305 different gap-filled measurements used to compute a half-hourly flux. The largest contribution
 306 to gap-filled measurements was soil water. SJER and WOOD utilized the largest number of
 307 gap-filled measurements, which were primarily SWC and T_S .



308 Figure 6: Panel a) Proportion of input gap-filled environmental measurements used to generate
 309 F_S from the `neonSoilFlux` package, by study site. Panel b) distribution of the usage
 310 of gap-filled measurements at each site.

308 Figure 7 reports both the computed SNR and the proportion of measured field fluxes within the
 309 modeled uncertainty for a given flux computation method F_{ijk} (Section 4.3). Here, values of

Table 2

	Millington-Quirk		Marshall	
	NRMSE	R2	NRMSE	R2
KONZ				
F_{110}	0.87	0.41	0.63	0.41
F_{101}	0.69	0.22	0.60	0.15
F_{011}	0.52	0.20	1.35	0.25
F_{000}	0.70	0.23	0.58	0.14
SJER				
F_{110}	0.13	0.17	0.14	0.19
F_{101}	0.32	0.21	0.31	0.24
F_{011}	0.49	0.02	0.48	0.03
F_{000}	0.29	0.18	0.28	0.19
SRER				
F_{110}	0.56	0.00	0.59	0.00
F_{101}	0.66	0.53	0.67	0.52
F_{011}	0.69	0.49	0.70	0.49
F_{000}	0.58	0.51	0.61	0.51
UNDE				
F_{110}	0.76	0.10	0.25	0.02
F_{101}	0.97	0.28	0.66	0.21
F_{011}	0.97	0.15	0.66	0.06
F_{000}	0.70	0.30	0.38	0.05
WOOD				
F_{110}	0.44	0.03	0.93	0.02
F_{101}	0.89	0.07	0.74	0.05
F_{011}	1.12	0.02	1.22	0.01
F_{000}	0.56	0.06	0.46	0.05
WREF				
F_{110}	0.53	0.78	0.35	0.75
F_{101}	0.91	0.24	0.73	0.35
F_{011}	1.03	0.37	1.07	0.37
F_{000}	0.84	0.00	0.49	0.05

310 SNR greater than unity indicates a reported uncertainty is smaller, propagated by quadrature
 311 from a relatively higher precision from measured input variables (CO_2 , T_S , SWC , or P). The
 312 sensitivity to the uncertainty reduction factor (ϵ , bottom panels in Figure 7) demonstrates
 313 how accuracy could be improved if modeled uncertainty σ_{ijk} decreases.

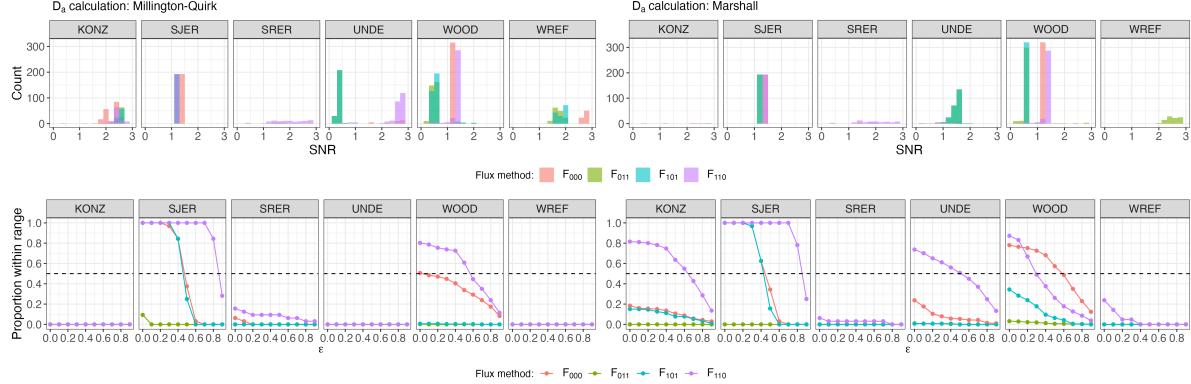


Figure 7: Top panels: distribution of SNR values across each of the different sites for modeled effluxes from the `neonSoilFlux` package, depending on the diffusivity calculation used (Millington-Quirk or Marshall, Section 4.2.2). Bottom panels: Proportion of measured F_S within the modeled range of a flux computation method F_{ijk} given an uncertainty reduction factor ϵ , or $|F_S - F_{ijk}| < (1 - \epsilon)\sigma_{ijk}$.

314 Figure 8 reports the distribution of D_a (from both the Marshall and Millington-Quirk methods,
 315 Section 4.2.2) at each study site, and the *post hoc* computation of D_a (Section 4.2.2).

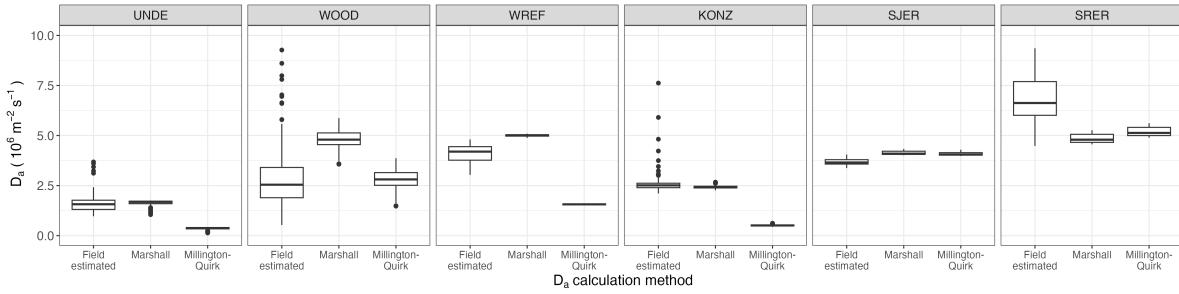


Figure 8

316 **6 Discussion**

317 This study presents a unified data science workflow to efficiently process automated measure-
318 ments of belowground soil CO₂ concentrations, water, and temperature to infer estimates of
319 soil surface CO₂ effluxes through application of Fick's Law (Equation 4). Our core goals in this
320 study were: (1) to generate estimates of soil flux from continuous soil sensor data at terrestrial
321 NEON sites using the flux-gradient method and then (2) to compare those estimates to field-
322 measured fluxes based on the closed chamber approach at six NEON focal sites. We discuss
323 our progress toward these core goals through (1) an overall evaluation of the flux-gradient ap-
324 proach (and uncertainty calculation) and (2) site-specific evaluation of differences in estimated
325 vs measured fluxes.

326 **6.1 General evaluation of flux-gradient approach**

327 Key assumptions of the flux-gradient approach are that CO₂ concentrations increase through-
328 out the soil profile. We found that this condition was met at XXX% across the study period.
329 Periods where this gradient condition are not met generally are connected to biophysical
330 processes such soil wetting events (e.g. KONZ), which have the effect of reducing the soil res-
331piration or efflux due to a temporary reduction in diffusivity. When modeling soil respiration,
332 typically a non-linear response function that also considers soil type is used (Bouma & Bryla,
333 2000; Yan et al., 2016, 2018). For the `neonSoilFlux` package, soil type is connected to the bulk
334 density, which was characterized at each NEON site based on replicate samples collected from
335 the site megapit at a subset of soil horizons, with an estimated uncertainty of ±5% (see NEON
336 User Guide to Soil physical and chemical properties, Megapit (DP1.00096.001)). Coarse frag-
337ment estimates also have very large uncertainties, but because the volume fraction tends to
338 be low in surface soils it probably wouldn't contribute much additional flux uncertainty.

339 The largest source of uncertainty to improve reliability of the flux estimate is to prevent the
340 usage of gap-filled data. Three sites (KONZ, SRER, and KONZ) had more than 75% of half-
341 hourly periods with no-gap filled measurements. Two sites (SJER and WOOD) had more
342 than 75% of half-hourly intervals with just one gap-filled measurement. While WREF re-
343 ported no gap-filled measurements, field data collection occurred following a once-in-a century
344 rainstorm with soils observed at their water holding capacity. We recommend that whenever
345 available, local field knowledge is supplementary to any QA filtering protocol of fluxes from
346 the `neonSoilFlux` package.

347 We recognize that this gap-filling approach may lead to gap-filled values that are quite different
348 from the actual values, such as an underestimate of soil moisture following rain events. Further
349 extensions of the gap filling method could use more sophisticated gap-filling routines, similar to
350 what is used for net ecosystem carbon exchange (Falge et al., 2001; Liu et al., 2023; Mariethoz
351 et al., 2015; Moffat et al., 2007; Zhang et al., 2023). The current gap-filling routine provides
352 a consistent approach that can be applied to each data stream, but further work may explore
353 alternative gap-filling approaches.

354 Based on this approach, we would *a priori* expect $F_{011} \leq F_{101} \leq F_{110} \leq F_{000}$ because the
355 previous flux estimates ones correspond to deeper depths which will could miss CO₂ produced
356 in shallower layers. Additionally, field flux measurements should correlate with F_{000} because
357 they represent surface fluxes.

358 **6.2 Evaluation of flux-gradient approach at each site**

359 Derived results from the `neonSoilFlux` package have patterns that are consistent, and
360 comparable, to those directly measured to the field (Figure XXX). The advantage to the
361 `neonSoilFlux` package is the calculation of fluxes across different measurement depths,

allowing for additional site-specific customization. Here application of the flux-gradient method provides a baseline estimate of soil fluxes that could be complemented through additional field measurements (e.g. LICOR).

The six sites studied provide separate case studies for considerations when applying the flux-gradient method to evaluate resulting uncertainties and fluxes. For example, SRER is characterized by sandy soil, which also led to the highest observed field soil temperatures. At SRER the flux across the top two layers (F_{110}) produced a pattern of soil flux consistent with the observed field data. The remaining methods F_{101} , F_{011} , or F_{000} are derived from information at the deeper layer, which is decoupled both in terms of temperature and CO₂ concentration.

In addition, KONZ is a site that experienced a significant rain event prior to sampling with eventual drying out over the course of the experiment. In this case we observed storage of soil water which increased the soil CO₂ at the top layer, leading to negative values of flux at the start of the experiment, with the fluxes drying out afterwards. In this case only when the soil dried out (or returned to a baseline level), that the fluxes at the provided layer would work out in this case.

When considering systematic deployment of this method across a measurement network, we faced a number of independent challenges for consideration.

Figure 7 illustrates the tradeoff between accuracy for modeled fluxes (defined here as closeness to field-measured F_S) and precision defined by the SNR, and how this is confounded by the choice of diffusivity model used. MORE HERE

Diffusivity discussion

In developing and validating our approach, we faced a number of challenges related to data availability, including... gap filling, sensor calibration, depth interpolation, rainstorms, etc

385 These errors are all

386 **6.3 Recommendations for future method development**

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

393 Figures XXX suggests that the provided uncertainty from `neonSoilFlux` is an overestimate
394 compared to what is actually computed. When $\epsilon = 0$ in Figure Figure 7, that means we
395 are just using the reported uncertainty from `neonSoilFlux`. Looking at that ($\text{epsilon} = 0$)
396 shows field measurements UNDE, KONZ, SJER are 100% within the reported intervals from
397 `neonSoilFlux`. But those sites tend to have a $\text{SNR} < 1$, so the uncertainty is pretty noisy. For
398 UNDE, we could even reduce the uncertainty by a factor of 75% ($\text{epsilon} = 0.75$), more than
399 half of the field measurements will still be within the reported intervals. For KONZ, we are
400 still within 70% of the reported intervals when uncertainty is reduced by 90%. That suggests
401 that while the reported accuracy (as compared to field measurements), we do have higher
402 precision.

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

⁴⁰⁷ turn, aids in our ability to understand the components of net ecosystem flux assessed at these
⁴⁰⁸ sites using the co-located eddy flux towers.

- ⁴⁰⁹ • Refine estimates to provide a realistic constraint on surface concentration measurements,
⁴¹⁰ thereby increasing the gradient.
⁴¹¹ • Apply machine learning algorithms (e.g. random trees) or model averaging techniques to
⁴¹² generate a single flux estimate across each sites spatial location
⁴¹³ • Benchmarking flux results to estimates provided by Net ecosystem carbon exchange.

⁴¹⁴ **7 Conclusions**

⁴¹⁵ We have here presented an R package `neonSoilFlux` for the estimation of soil CO₂ fluxes from
⁴¹⁶ continuous buried soil sensor measurements across terrestrial National Ecological Observatory
⁴¹⁷ Network sites. We compared the predicted fluxes to those measured directly using a field-based
⁴¹⁸ closed chamber approach. We find that...

- ⁴¹⁹ Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the
⁴²⁰ atmosphere - the state and future of the eddy covariance method. *Global Change Biology*,
⁴²¹ 20(12), 3600–3609. <https://doi.org/10.1111/gcb.12649>
- ⁴²² Berenbaum, M. R., Carpenter, S. R., Hampton, S. E., Running, S. W., & Stanzione, D. C.
⁴²³ (2015). *Report from the NSF BIO Advisory Committee Subcommittee on NEON Scope*
⁴²⁴ *Impacts*.
- ⁴²⁵ Bond-Lamberty, B. (2018). New Techniques and Data for Understanding the Global Soil Res-
⁴²⁶piration Flux. *Earth's Future*, 6(9), 1176–1180. <https://doi.org/10.1029/2018EF000866>
- ⁴²⁷ Bond-Lamberty, B., Ballantyne, A., Berryman, E., Fluet-Chouinard, E., Jian, J., Morris, K.
⁴²⁸ A., Rey, A., & Vargas, R. (2024). Twenty Years of Progress, Challenges, and Opportuni-

- ties in Measuring and Understanding Soil Respiration. *Journal of Geophysical Research: Biogeosciences*, 129(2), e2023JG007637. <https://doi.org/10.1029/2023JG007637>
- Bond-Lamberty, B., Christianson, D. S., Malhotra, A., Pennington, S. C., Sihi, D., AghaKouchak, A., Anjileli, H., Altaf Arain, M., Armesto, J. J., Ashraf, S., Ataka, M., Baldocchi, D., Andrew Black, T., Buchmann, N., Carbone, M. S., Chang, S.-C., Crill, P., Curtis, P. S., Davidson, E. A., ... Zou, J. (2020). COSORE: A community database for continuous soil respiration and other soil-atmosphere greenhouse gas flux data. *Global Change Biology*, 26(12), 7268–7283. <https://doi.org/10.1111/gcb.15353>
- Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Biogeosciences*, 7(6), 1915–1926. <https://doi.org/10.5194/bg-7-1915-2010>
- Bond-Lamberty, B., Wang, C., & Gower, S. T. (2004). A global relationship between the heterotrophic and autotrophic components of soil respiration? *Global Change Biology*, 10(10), 1756–1766. <https://doi.org/10.1111/j.1365-2486.2004.00816.x>
- Bouma, T. J., & Bryla, D. R. (2000). On the assessment of root and soil respiration for soils of different textures: Interactions with soil moisture contents and soil CO₂ concentrations. *Plant and Soil*, 227(1), 215–221. <https://doi.org/10.1023/A:1026502414977>
- Chen, H., & Tian, H.-Q. (2005). Does a General Temperature-Dependent Q10 Model of Soil Respiration Exist at Biome and Global Scale? *Journal of Integrative Plant Biology*, 47(11), 1288–1302. <https://doi.org/10.1111/j.1744-7909.2005.00211.x>
- Davidson, E. A., Janssens, I. A., & Luo, Y. (2006). On the variability of respiration in terrestrial ecosystems: Moving beyond Q10. *Global Change Biology*, 12, 154–164. <https://doi.org/10.1111/j.1365-2486.2005.01065.x>
- Desai, A. R., Murphy, B. A., Wiesner, S., Thom, J., Butterworth, B. J., Koupaei-Abyazani, N., Muttaqin, A., Paleri, S., Talib, A., Turner, J., Mineau, J., Merrelli, A., Stoy, P., & Davis, K. (2022). Drivers of Decadal Carbon Fluxes Across Temperate Ecosystems. *Journal of Geophysical Research: Biogeosciences*, 127(12), e2022JG007014. <https://doi.org/10.1029/2022JG007014>

- 455 2022JG007014
- 456 Elshall, A. S., Ye, M., Pei, Y., Zhang, F., Niu, G.-Y., & Barron-Gafford, G. A. (2018). Relative
457 model score: A scoring rule for evaluating ensemble simulations with application to micro-
458 microbial soil respiration modeling. *Stochastic Environmental Research and Risk Assessment*,
459 32(10), 2809–2819. <https://doi.org/10.1007/s00477-018-1592-3>
- 460 Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G.,
461 Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger,
462 D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., ... Wofsy, S. (2001).
463 Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural
464 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)
- 465 Farrance, I., & Frenkel, R. (2012). *Uncertainty of Measurement: A Review of the Rules
466 for Calculating Uncertainty Components through Functional Relationships*. *The Clinical
467 Biochemist Reviews*, 33(2), 49–75.
- 468 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck,
469 J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J.,
470 Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni,
471 P., ... Zheng, B. (2023). Global Carbon Budget 2023. *Earth System Science Data*, 15(12),
472 5301–5369. <https://doi.org/10.5194/essd-15-5301-2023>
- 473 Hamdi, S., Moyano, F., Sall, S., Bernoux, M., & Chevallier, T. (2013). Synthesis analysis
474 of the temperature sensitivity of soil respiration from laboratory studies in relation to
475 incubation methods and soil conditions. *Soil Biology and Biochemistry*, 58, 115–126. <https://doi.org/10.1016/j.soilbio.2012.11.012>
- 476 Hirano, T., Kim, H., & Tanaka, Y. (2003). Long-term half-hourly measurement of soil CO₂
477 concentration and soil respiration in a temperate deciduous forest. *Journal of Geophysical
478 Research: Atmospheres*, 108(D20). <https://doi.org/10.1029/2003JD003766>
- 479 Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017).

- 481 The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls.
482 *Annual Review of Ecology, Evolution and Systematics*, 48(Volume 48, 2017), 419–445.
483 <https://doi.org/10.1146/annurev-ecolsys-112414-054234>
- 484 Jian, J., Bailey, V., Dorheim, K., Konings, A. G., Hao, D., Shiklomanov, A. N., Snyder, A.,
485 Steele, M., Teramoto, M., Vargas, R., & Bond-Lamberty, B. (2022). Historically inconsis-
486 tent productivity and respiration fluxes in the global terrestrial carbon cycle. *Nature
487 Communications*, 13(1), 1733. <https://doi.org/10.1038/s41467-022-29391-5>
- 488 Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N.,
489 Manzon, J., Marchesi, R., Paredes, D., & Bond-Lamberty, B. (2021). A restructured and
490 updated global soil respiration database (SRDB-V5). *Earth System Science Data*, 13(2),
491 255–267. <https://doi.org/10.5194/essd-13-255-2021>
- 492 Jiang, J., Feng, L., Hu, J., Liu, H., Zhu, C., Chen, B., & Chen, T. (2024). Global soil
493 respiration predictions with associated uncertainties from different spatio-temporal data
494 subsets. *Ecological Informatics*, 82, 102777. <https://doi.org/10.1016/j.ecoinf.2024.102777>
- 495 Jobbágy, E. G., & Jackson, R. B. (2000). The Vertical Distribution of Soil Organic Carbon
496 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)
- 497 Liu, K., Li, X., Wang, S., & Zhang, H. (2023). A robust gap-filling approach for European
498 Space Agency Climate Change Initiative (ESA CCI) soil moisture integrating satellite
499 observations, model-driven knowledge, and spatiotemporal machine learning. *Hydrology
500 and Earth System Sciences*, 27(2), 577–598. <https://doi.org/10.5194/hess-27-577-2023>
- 502 Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011).
503 Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*,
504 21(5), 1429–1442. <https://doi.org/10.1890/09-1275.1>
- 505 Maier, M., & Schack-Kirchner, H. (2014). Using the gradient method to determine soil gas
506 flux: A review. *Agricultural and Forest Meteorology*, 192–193, 78–95. <https://doi.org/10.>

- 507 1016/j.agrformet.2014.03.006
- 508 Mariethoz, G., Linde, N., Jougnot, D., & Rezaee, H. (2015). Feature-preserving interpolation
509 and filtering of environmental time series. *Environmental Modelling & Software*, 72, 71–76.
- 510 <https://doi.org/10.1016/j.envsoft.2015.07.001>
- 511 Marshall, T. J. (1959). The Diffusion of Gases Through Porous Media. *Journal of Soil Science*,
512 10(1), 79–82. <https://doi.org/10.1111/j.1365-2389.1959.tb00667.x>
- 513 Millington, R. J., & Shearer, R. C. (1971). Diffusion in aggregated porous media. *Soil Science*,
514 111(6), 372–378.
- 515 Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G.,
516 Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann,
517 M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., & Stauch, V. J. (2007). Comprehensive
518 comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and
Forest Meteorology*, 147(3), 209–232. <https://doi.org/10.1016/j.agrformet.2007.08.011>
- 519 Moldrup, P., Olesen, T., Yamaguchi, T., Schjønning, P., & Rolston, D. E. (1999). Modeling
520 diffusion and reaction in soils: 9. The Buckingham-Burdine-Campbell equation for gas
521 diffusivity in undisturbed soil. *Soil Science*, 164(2), 75.
- 522 National Ecological Observatory Network (NEON). (2024a). *Barometric pressure*
523 (*DP1.00004.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/RT4V-KZ04>
- 524 National Ecological Observatory Network (NEON). (2024b). *Soil CO₂ concentra-*
525 *tion* (*DP1.00095.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/E7GR-6G94>
- 526 National Ecological Observatory Network (NEON). (2024c). *Soil physical and chemical proper-*
527 *ties, Megapit* (*DP1.00096.001*). National Ecological Observatory Network (NEON). <https://doi.org/10.48443/S6ND-Q840>
- 528 National Ecological Observatory Network (NEON). (2024d). *Soil temperature* (*DP1.00041.001*).

- 533 National Ecological Observatory Network (NEON). <https://doi.org/10.48443/Q24X-PW21>
- 534 National Ecological Observatory Network (NEON). (2024e). *Soil water content and water*
535 *salinity (DP1.00094.001)*. National Ecological Observatory Network (NEON). <https://doi.org/10.48443/A8VY-Y813>
- 536
- 537 Norman, J. M., Kucharik, C. J., Gower, S. T., Baldocchi, D. D., Crill, P. M., Rayment, M.,
538 Savage, K., & Striegl, R. G. (1997). A comparison of six methods for measuring soil-
539 surface carbon dioxide fluxes. *Journal of Geophysical Research: Atmospheres*, 102(D24),
540 28771–28777. <https://doi.org/10.1029/97JD01440>
- 541 Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown,
542 K., & Vargas, R. (2017). The value of soil respiration measurements for interpreting and
543 modeling terrestrial carbon cycling. *Plant and Soil*, 413(1), 1–25. <https://doi.org/10.1007/s11104-016-3084-x>
- 544
- 545 Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). *Using Bayesian Model*
546 *Averaging to Calibrate Forecast Ensembles*. <https://doi.org/10.1175/MWR2906.1>
- 547 Sallam, A., Jury, W. A., & Letey, J. (1984). Measurement of Gas Diffusion Coefficient under
548 Relatively Low Air-filled Porosity. *Soil Science Society of America Journal*, 48(1), 3–6.
549 <https://doi.org/10.2136/sssaj1984.03615995004800010001x>
- 550 Shao, J., Zhou, X., Luo, Y., Li, B., Aurela, M., Billesbach, D., Blanken, P. D., Bracho, R.,
551 Chen, J., Fischer, M., Fu, Y., Gu, L., Han, S., He, Y., Kolb, T., Li, Y., Nagy, Z., Niu, S.,
552 Oechel, W. C., ... Zhang, J. (2015). Biotic and climatic controls on interannual variability
553 in carbon fluxes across terrestrial ecosystems. *Agricultural and Forest Meteorology*, 205,
554 11–22. <https://doi.org/10.1016/j.agrformet.2015.02.007>
- 555 Shao, P., Zeng, X., Moore, D. J. P., & Zeng, X. (2013). Soil microbial respiration from
556 observations and Earth System Models. *Environmental Research Letters*, 8(3), 034034.
557 <https://doi.org/10.1088/1748-9326/8/3/034034>
- 558 Sihi, D., Gerber, S., Inglett, P. W., & Inglett, K. S. (2016). Comparing models of microbial–

- 559 substrate interactions and their response to warming. *Biogeosciences*, 13(6), 1733–1752.
- 560 <https://doi.org/10.5194/bg-13-1733-2016>
- 561 Tang, J., Baldocchi, D. D., Qi, Y., & Xu, L. (2003). Assessing soil CO₂ efflux using continuous
562 measurements of CO₂ profiles in soils with small solid-state sensors. *Agricultural and Forest
563 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)
- 564 Tang, J., Misson, L., Gershenson, A., Cheng, W., & Goldstein, A. H. (2005). Continuous
565 measurements of soil respiration with and without roots in a ponderosa pine plantation
566 in the Sierra Nevada Mountains. *Agricultural and Forest Meteorology*, 132(3), 212–227.
567 <https://doi.org/10.1016/j.agrformet.2005.07.011>
- 568 Taylor, J. R. (2022). *An Introduction to Error Analysis: The Study of Uncertainties in Physical
569 Measurements, Third Edition* (3rd ed.). University Science Press.
- 570 Yan, Z., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S., Liu, C., & Liu, C. (2018).
571 A moisture function of soil heterotrophic respiration that incorporates microscale processes.
572 *Nature Communications*, 9(1), 2562. <https://doi.org/10.1038/s41467-018-04971-6>
- 573 Yan, Z., Liu, C., Todd-Brown, K. E., Liu, Y., Bond-Lamberty, B., & Bailey, V. L. (2016).
574 Pore-scale investigation on the response of heterotrophic respiration to moisture conditions
575 in heterogeneous soils. *Biogeochemistry*, 131(1), 121–134. [https://doi.org/10.1007/s10533-016-0270-0](https://doi.org/10.1007/s10533-
576 016-0270-0)
- 577 Zhang, R., Kim, S., Kim, H., Fang, B., Sharma, A., & Lakshmi, V. (2023). Temporal
578 Gap-Filling of 12-Hourly SMAP Soil Moisture Over the CONUS Using Water Balance
579 Budgeting. *Water Resources Research*, 59(12), e2023WR034457. [https://doi.org/10.1029/2023WR034457](https://doi.org/10.1029/
580 2023WR034457)