Supplemental information for `neonSoilFlux`: An R Package for Continuous Sensor-Based Estimation of Soil CO$_2$ Fluxes

1 Assessment of data gaps

For a given half-hourly time period, the `neonSoilFlux` packages assigns a QA flag for a measurement if more one values across all measurement depths uses gap-filled data (Section 4.2.1 of the main text). Panel a of Figure S1 reports the proportion of gap-filled data for all input environmental measurements at each site during the period when field measurements were made. Soil fluxes are computed from 4 different types of input measurements (T_S, SWC, P, and CO$_2$), any of which could have a QA flag in a half-hourly interval. Panel b of Figure S1 displays at each site the distribution of the number of different gap-filled measurements used to compute a half-hourly flux. The largest cause of measurements needing to be gap-filled was missing or flagged soil moisture data. Calculating fluxes for WOOD and SJER required using the largest proportion of gap-filled measurements, due to substantially large fractions of flagged or missing SWC and T_S data.
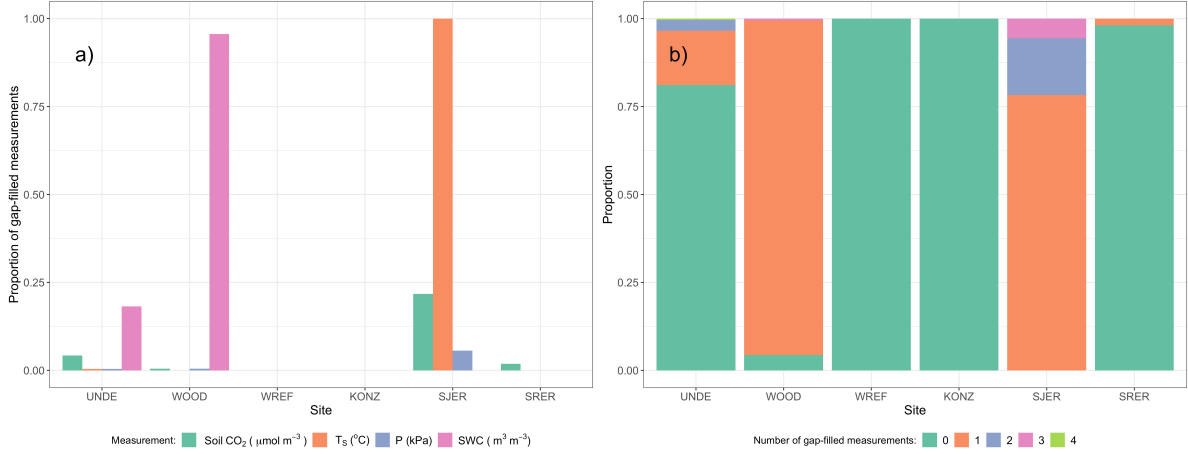
Figure S1: Panel a) Proportion of input gap-filled environmental measurements used to generate F_S from the `neonSoilFlux` package, by study site. Panel b) distribution of the usage of gap-filled measurements at each site.

2 Assessing the signal to noise ratio (SNR) and evaluating estimated uncertainties

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. Beyond the model statistics defined in the main text, we computed the signal to noise ratio (SNR), defined as the ratio of a modeled soil flux (F_{ijk}) from `neonSoilFlux` to its quadrature uncertainty (σ_{ijk}).

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

2

₃₁ The computed signal to noise ratio (SNR) and the proportion of measured field fluxes within
₃₂ the modeled uncertainty for a given flux computation method F_{ijk} suggest that there was
₃₃ substantial variability in the agreement between the gradient method and field-measured ob-
₃₄ servations (Figure S2, Section 4.3 of the main text). Here, values of SNR greater than unity
₃₅ indicate lower reported uncertainty, as propagated by quadrature due to a relatively higher
₃₆ precision of measured input variables (CO_2, T_S, SWC, or P).

₃₇ The sensitivity to an uncertainty reduction factor (ϵ, bottom panels in Figure S2) demonstrates
₃₈ how concordance between measured and modeled fluxes would be affected if environmental
₃₉ measurement uncertainty σ_{ijk} were to decrease. As ϵ increases from left to right in each figure,
₄₀ the possible range of values for each predicted flux value decreases and the proportion of
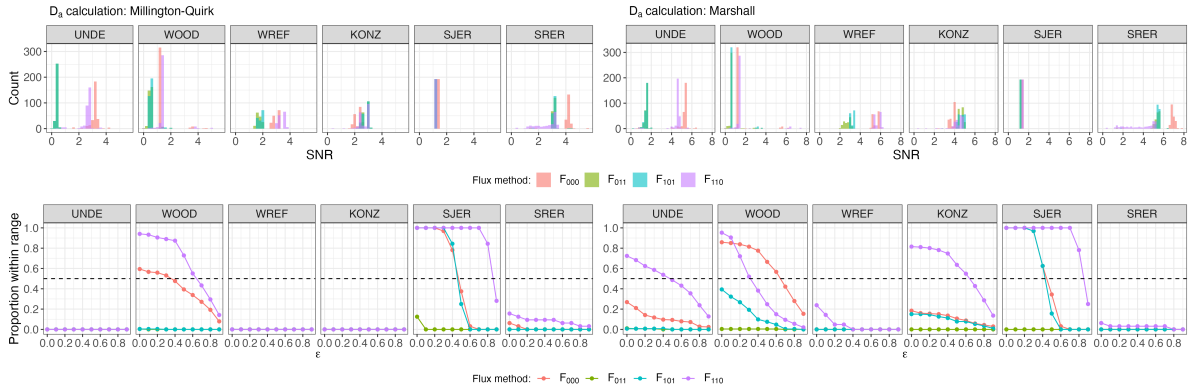₄₁ measured fluxes that fall within that range also decreases.

Figure S2: 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 of the main text). 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 $|F_S - F_{ijk}| < (1 - \epsilon)\sigma_{ijk}$.