JGR Biogeosciences

METHOD
10.1029/2023JG007451

Fluxbots: A Method for Building, Deploying, Collecting and Analyzing Data From an Array of Inexpensive, Autonomous Soil Carbon Flux Chambers

Elizabeth Forbes[1,2], Vincent Benenati[3,4], Spencer Frey[5], Mare Hirsch[6], George Koech[7], Grace Lewin[1,8], John Naisikie Mantas[7], and Kelly Caylor[8,9]

[1]Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA, USA, [2]School of the Environment, Yale University, New Haven, CT, USA, [3]Department of Computer Science, University of California Santa Barbara, Santa Barbara, CA, USA, [4]Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA, USA, [5]Department of Physics, University of California Santa Barbara, Santa Barbara, CA, USA, [6]Media Arts and Technology Graduate Program, University of California Santa Barbara, Santa Barbara, CA, USA, [7]Mpala Research Centre and Conservancy, Laikipia County, Kenya, [8]Bren School of Environmental Science and Management, University of California Santa Barbara, Santa Barbara, CA, USA, [9]Department of Geography, University of California Santa Barbara, Santa Barbara, CA, USA

Abstract Soil carbon flux rates are a crucial metric of carbon cycling that contribute to calculating an ecosystem's carbon budget, and thus whether it is a source or sink of atmospheric carbon dioxide. However, soil carbon flux datasets are frequently low-resolution across either space or time, limiting our abilities to identify small-scale ecological contexts that influence soil carbon dynamics. Existing datasets are distributed unevenly, with some soil carbon-rich regions (like tropical grasslands) significantly understudied. We developed an autonomous, inexpensive, do-it-yourself (DIY) soil carbon flux chamber (a "fluxbot") and data processing software. We deployed a distributed array of 12 fluxbots in a long-term experiment in a central Kenyan savanna where it has been logistically impossible to collect high-resolution soil carbon flux data. With this array we collected over 10,000 individual flux estimates over almost two months, spanning the end of a dry season and the start of a wet season. With our successful deployment in situ, we demonstrate the potential for low-cost, autonomous, DIY sensors in improving resolution of soil carbon flux datasets (particularly in under-studied or logistically challenging systems). If implemented widely, such an improvement in data collection capacities could improve our understanding of ecological and climatic drivers of soil carbon flux dynamics on the local to global scale.

Plain Language Summary Soil carbon flux, the rate at which carbon dioxide is exchanged between soil and the atmosphere, is a key feature of an ecosystem's carbon budget. However, measuring soil carbon flux rates at spatial and temporal scales that capture global ecosystems' ecological heterogeneity is extremely difficult due to the logistical constraints of manual data collection and high costs of commercial sensor systems. As such, many existing soil carbon flux datasets do not have the resolution necessary to identify small-scale ecological patterns in carbon dynamics, and datasets are distributed unevenly across easy-to-monitor ecosystems globally. We developed an inexpensive, robotic, autonomous soil carbon flux chamber that collects hourly data for as long as it is deployed. We built and deployed an array of 12 sensors in an ecologically complex central Kenyan savanna ecosystem. We collected almost two months of hourly data consisting of over 10,000 soil carbon flux measurements, the largest and most high-resolution dataset collected in this system. Wider adoption of such open-access chambers could result in the collection of highly resolved soil carbon flux datasets in understudied systems worldwide, and greater understanding of the ecological contexts that mediate soil carbon flux.

Key Points:

- We developed an autonomous soil carbon flux chamber ("fluxbot") to improve data collection resolution and extent
- We prioritized low price and accessibility to enable the capture of small-scale ecological heterogeneity
- We distributed a fluxbot array in a large community ecology experiment to test their ability to capture complexity

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. Forbes,
forbeselizabeth.s@gmail.com

Citation:

Forbes, E., Benenati, V., Frey, S., Hirsch, M., Koech, G., Lewin, G., et al. (2023). Fluxbots: A method for building, deploying, collecting and analyzing data from an array of inexpensive, autonomous soil carbon flux chambers. *Journal of Geophysical Research: Biogeosciences*, *128*, e2023JG007451. https://doi.org/10.1029/2023JG007451

Received 27 FEB 2023
Accepted 23 MAY 2023

Author Contributions:

Conceptualization: Elizabeth Forbes, Kelly Caylor
Data curation: Elizabeth Forbes, Kelly Caylor
Formal analysis: Elizabeth Forbes, Kelly Caylor
Funding acquisition: Elizabeth Forbes, Kelly Caylor
Investigation: Elizabeth Forbes, Vincent Benenati, Spencer Frey, Mare Hirsch, George Koech, Grace Lewin, John Naisikie Mantas, Kelly Caylor

1. Introduction

As the Anthropocene continues, predicting the extent of climate change impacts is increasingly crucial. The biogeochemical research community has a long, successful history of monitoring ecosystem carbon budgets (i.e., the balance between the atmospheric carbon they take up and what they emit) on local to global scales. This kind of monitoring has historically been accomplished with observatory sites all over the world; for example, eddy flux towers that use the turbulence of air movement to extrapolate fluxes of atmospheric gases and energetic

Methodology: Elizabeth Forbes, Spencer Frey, Mare Hirsch, Grace Lewin, Kelly Caylor
Project Administration: Elizabeth Forbes, George Koech, John Naisikie Mantas
Resources: Elizabeth Forbes, Kelly Caylor
Software: Elizabeth Forbes, Vincent Benenati, Mare Hirsch, Kelly Caylor
Supervision: Elizabeth Forbes
Validation: Elizabeth Forbes, Vincent Benenati, Spencer Frey, Kelly Caylor
Visualization: Elizabeth Forbes
Writing – original draft: Elizabeth Forbes, Kelly Caylor
Writing – review & editing: Elizabeth Forbes, Vincent Benenati, Spencer Frey, Mare Hirsch, George Koech, Grace Lewin, John Naisikie Mantas, Kelly Caylor

properties from an ecosystem. Today, there is a globally-distributed array of such towers monitoring fluxes, including carbon dioxide (CO_2) (Baldocchi et al., 2001). These data allow researchers to calculate a landscape's CO_2 emissions, and parameterize models with which to predict future emissions scenarios (Xiao et al., 2012).

In such terrestrial systems, one of the fluxes that makes up a carbon budget (and that can be used to validate ecosystem-scale carbon cycling models [Keenan et al., 2012; Phillips et al., 2017]) is soil carbon flux: the rate at which CO_2 is exchanged between the soil and the atmosphere. Soil carbon flux consists of biotic soil respiration (the combined respiration of microfauna, macrofauna, and roots within the soil), and is moderated by air and soil temperature, atmospheric pressure, soil moisture, soil type, and soil morphology (DeCarlo & Caylor, 2020). On a global scale, soils contain three times as much carbon as either the atmosphere or vegetation (Schmidt et al., 2011). A high-quality understanding of patterns in soil carbon flux is therefore essential to understanding the drivers of landscape-scale terrestrial CO_2 emissions.

On large scales, soil carbon flux is influenced by short- and long-term climatic processes, like individual weather events and seasonality (Delgado-Baquerizo et al., 2017; Munson et al., 2010) and by landscape-scale changes to ecosystem structure driven by fire (Pellegrini et al., 2020), land use (Wachiye et al., 2020), and agriculture (Lohila et al., 2003; Rochette et al., 1991). Critically, it is also influenced by extremely small-scale variability (Rodeghiero & Cescatti, 2008), like patches of vegetation (Stoyan et al., 2000), wind or water deposition of organic matter (Throop et al., 2020), and geomorphic idiosyncrasies like the many subsurface cracks, chimneys and pipes that form (Holden, 2005) or are engineered by large (DeCarlo & Caylor, 2019) and small (Ohashi et al., 2007; Otieno et al., 2011) fauna. Considering that soil carbon flux can make up between 30% and 80% of net ecosystem exchange (e.g., Davidson et al. (2006)), it is important to parse how distinct sources of small-scale ecological heterogeneity contribute to an ecosystem's overall carbon cycling dynamics.

The high-frequency (greater than 24 hr), high resolution (less than 1 km^2) datasets collected at eddy flux towers are critical to answering hydrological, earth system, and climatological questions (Baldocchi et al., 2001). However, this resolution remains too coarse for identifying the many small-scale, within-landscape drivers of soil carbon flux, and these datasets are less easily interrogated from an ecological (particularly community ecology) point of view (Fisher & Koven, 2020; Tuovinen et al., 2019). Fine-scale flux data can instead be taken manually, using commercial in situ non-steady-state survey chambers placed over the soil that monitor the buildup of CO_2 over time (Davidson et al., 2002). Autonomously-operating commercial chambers can also be installed in situ, operating at a user-defined frequency. However, these systems are extremely expensive (approximately 25–50k USD, not including soil chamber attachments); are often multiplexed and limited in spatial distribution; and (in the case of survey chambers) require multiple operators and the purchase of multiple systems to measure flux at high resolution in space or time. In landscapes with challenging environmental conditions like seasonal rainfall or wildlife, use of these systems introduces risk to the operator, the expensive instruments, or both. Therefore despite their availability, there are substantial barriers to using commercial soil carbon flux chambers across landscape-scale extents or at high spatial or temporal resolution (Jian et al., 2021).

Ecologists can overcome this challenge with a do-it-yourself (DIY) approach, by building autonomous soil carbon flux chambers from low-cost materials and CO_2 sensors (Bastviken et al., 2015; Carbone et al., 2008, 2011; Harmon et al., 2015). By reducing the cost of individual autonomous chambers, researchers can deploy distributed arrays of chambers operating simultaneously, and therefore capture small-scale heterogeneity in soil carbon flux. Such highly resolved datasets would more fully characterize dynamism in soil carbon flux, and could be interrogated through a community ecology lens. While some degree of accuracy or precision in raw CO_2 detection may be sacrificed when using low-cost CO_2 sensors (compared to those in commercial options [Martin et al., 2017; Yasuda et al., 2012]), the capacity for collecting large scale, high resolution, ecologically meaningful data at replications high enough to approach a landscape's representative fluxes is gained in return (Adachi et al., 2005; Pantani et al., 2020).

Here we present a design and analysis plan for an autonomous, robotic, non-steady-state soil carbon flux sensing chamber (a "fluxbot"). We deployed an array of 12 fluxbots in a central Kenya savanna ecosystem, in which ecosystem structure is predictably homogeneous on a large scale but heterogeneous on centimeter to meter scales (DeCarlo & Caylor, 2020). We argue that for ecologists (particularly community ecologists who aim to relate patterns in soil carbon flux to an ecosystem's structure and community), prioritizing reproducibility and replication with low-cost sensing is an exciting and viable path forward.

2. Methods

2.1. Chamber Construction

We constructed each fluxbot with easy-to-find materials, purchased from online retailers, or hardware and electronics stores (Table S2). We fabricated the chamber of each from two lengths of $5 \times 5''$ square polyvinyl chloride (PVC) polymer pipe, cut to 7″ (chamber "body") and 2.5″ (chamber "lid"), connected at one side (the chamber's back) with hinges (Figure S1 in Supporting Information S1). (For more details on construction, including additions of gaskets and tests for air-tightness when closed [Figure S2 in Supporting Information S1], see Texts S1 and S2 in Supporting Information S1.)

We topped the lid with a custom-cut acrylic plate, using weather- and air-tight glue. We mounted the linear actuator to the body and lid via long horizontal screws, screwed from the outside to the inside on the left side of each. This mounting allows for free rotational movement on each end of the actuator as it extends (opening the lid to 90°) and contracts (sealing the chamber shut).

2.2. Hardware

Once the chambers were built, we installed a custom hardware system in each lid, constructed from a microcontroller (Pycom LoPy4, with Pycom Expansion Board), a calibrated (Text S3 in Supporting Information S1) non-dispersive infrared (NDIR) CO_2 sensor (CozIR LP miniature 5,000 ppm CO_2 sensor), a real-time clock (RTC; set to local time at our field site in Laikipia, Kenya) to ensure synchronized activity across all fluxbots, and a combined air pressure, temperature, and humidity sensor (Table S2). Data was written to a microSD card on the microcontroller.

We calibrated the CO_2 sensors with a single-point calibration procedure, by sealing each sensor inside the SRC-2 flux chamber attachment of a CIRAS-brand gas exchange system. We then set the calibration point to the CO_2 concentration detected inside the chamber by the CIRAS, after the concentration equilibrated and stabilized (see Text S3 in Supporting Information S1 for details).

We powered the electronics system with a V44 battery, which was charged with a 6W solar panel (Figure 1). We fixed all electronics in place to a custom acrylic mounting board (Figure S3 in Supporting Information S1), then fixed one of these "electronics systems" to the inside of each fluxbot lid from the top plate, with the CO_2 sensor facing the inside of the chamber volume, and connecting the linear actuator (installed with screws to the chamber body and lid) to the microcontroller (Figure 2). The total cost of materials for one fluxbot was $361.71 (US dollars) (Table S2), at least two orders of magnitude cheaper than the majority of commercially available options.

2.3. Software

The fluxbots run on a series of MicroPython scripts designed to support synchronous measurements across an array of fluxbots for the duration of their field installation (Forbes & Caylor, 2021). Upon powering on, the fluxbot's electronics system runs self-diagnostics; if a hardware error is detected, a light-emitting diode light blinks colors coded to specific errors until it is resolved (Text S4 and Table S1 in Supporting Information S1; visible in Figure 2). Upon successful diagnostics, the fluxbot begins its measurement schedule which loops every hour for as long as power is available. If power is lost temporarily, the loop begins again at the top of the next hour after power is restored. During the first 55 min of each hour, while open, the fluxbots collect ambient (CO_2 concentration, temperature, humidity, and atmospheric pressure) data at regular intervals (Figure 3). For the last 5 min of each hour, while closed, the fluxbots collect high-frequency data (once per second) data (Figure S4 in Supporting Information S1).

2.4. Field Installation, Data Collection

We deployed 12 fluxbots total in two treatment plots of a large-scale, long-term large herbivore exclusion experiment (the Kenya Long term Exclosure Experiment or KLEE), located at Mpala Research Centre and Conservancy in Laikipia, Kenya. Within each of a fenced, large-bodied herbivore exclosure plot and an unfenced plot where large herbivores remain, we installed three fluxbots on open soil patches and three beneath the canopies of

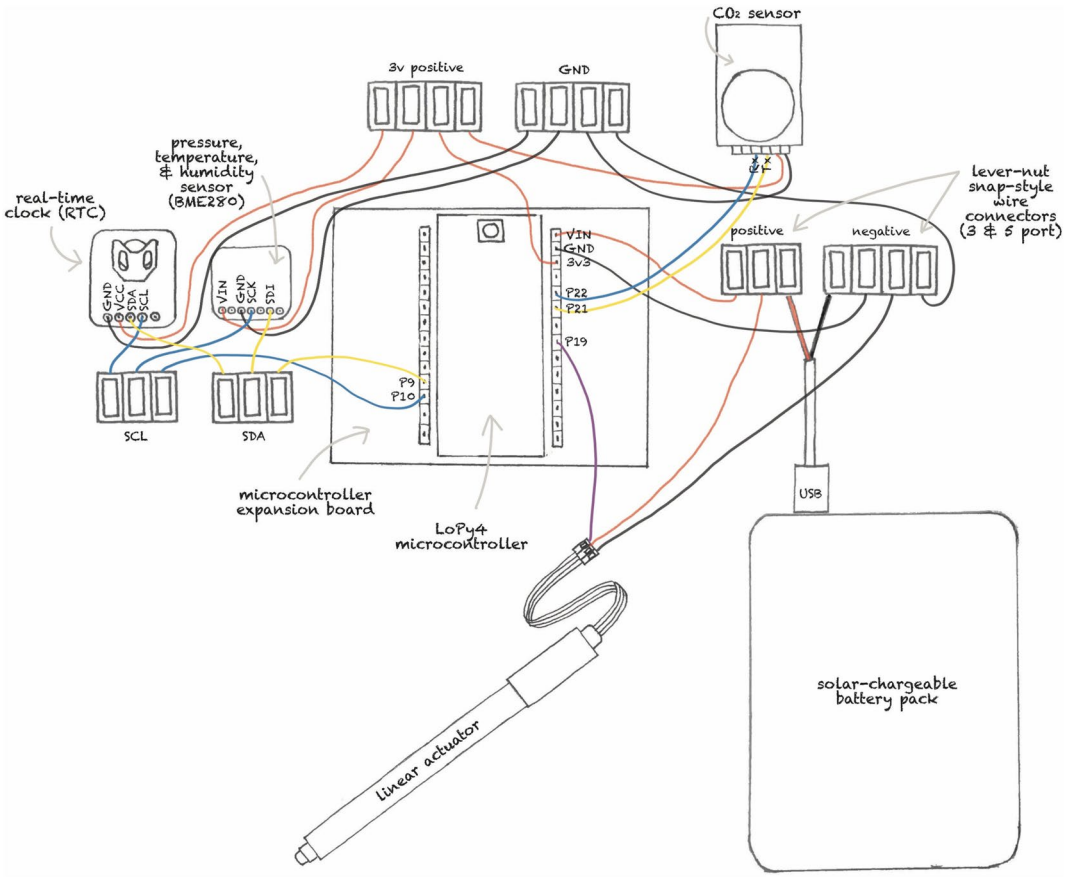

Figure 1. Wiring schematic, all components displayed on a plane.

the whistling thorn Acacia, *Acacia drepanolobium* (Figure S6 in Supporting Information S1). (*A. drepanolobium* have a mutualistic association with nitrogen-fixing bacteria in their root systems, and the soils beneath them are significantly enriched in bioavailable soil nitrogen [Fox-Dobbs et al., 2010].) We installed each fluxbot at least 10 m apart to ensure spatial independence. We fastened each fluxbot's solar panel horizontally to an additional

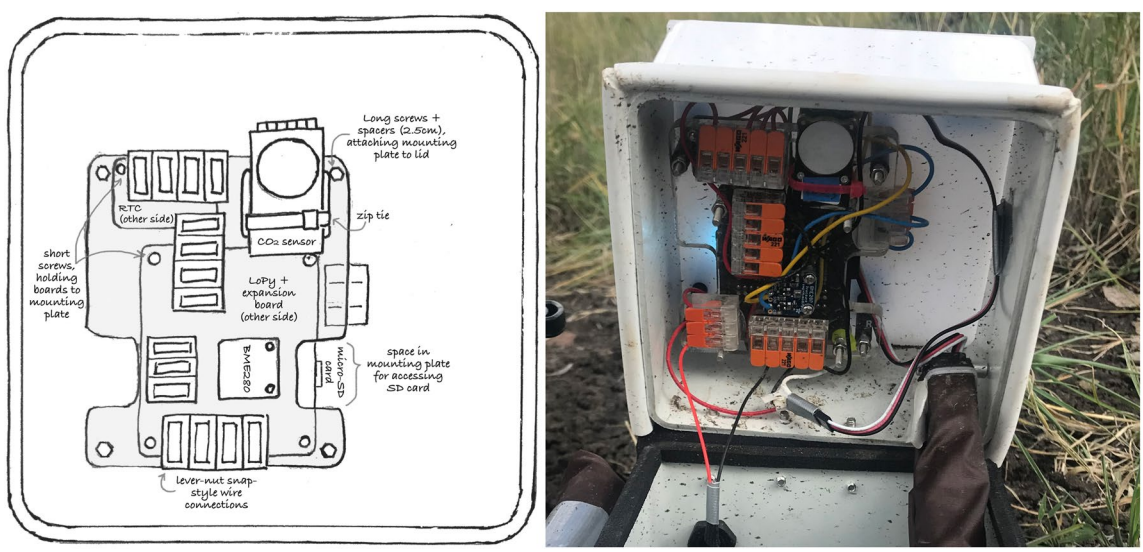

Figure 2. Electronics, installed, viewed inside the lid from the underside.

Figure 3. A fluxbot installed in the field, in its open position, with actuator fully extended. Collar for manual data collection at bottom right.

length of PVC to maximize sun exposure (Figure 3 and Figure S5 in Supporting Information S1).

The fluxbots remained in the field for approximately 2.5 months, over a period encompassing the end of a dry season and the start of a wet season (August 2019 to mid October 2019). We manually transferred the data from each fluxbot's microSD card to a Panasonic Toughbook laptop every 3–4 days, and copied the data to a hard drive and cloud storage service immediately upon return to the lab. The microSD cards are accessible when the fluxbots are open and removing it pauses fluxbot activity. The data download process takes several minutes at most; upon replacement of the microSD card each fluxbot boots up and resumes scheduled activity at the top of the next hour.

In addition to fluxbot data, in August 2019 we opportunistically collected soil carbon flux data manually from 12 round PVC collars installed adjacent to each fluxbot (visible in Figure 3 and Figure S5 in Supporting Information S1). We collected these data using a CIRAS gas exchange system with SRC-2 chamber attachment between 9 a.m. and 4 p.m., totaling just over 200 manually-collected flux measurements. We paired flux observations from co-located fluxbots and collars that occurred within 25 min of each other, and collated them to create a dataset of temporally- and spatially-associated fluxes (Text S7 in Supporting Information S1).

3. Flux Calculations

3.1. CO_2 Observations

The NDIR sensor contained in our fluxbots observes the quantity of CO_2 as a concentration in units of parts per million, $[CO_2]_{Obs}$, [ppm]. We convert $[CO_2]_{Obs}$ into a mass density of CO_2, presented here as the following series of conversions for ease of replication.

If we were observing the buildup of $[CO_2]_{Obs}$ in *dry* air, we would simply convert $[CO_2]_{Obs}$ to mass density by multiplying it by the molar density of dry air and the molar mass of CO_2. However, we account for the ratio of moist to dry air by calculating a humidity ratio to apply to the molar mass of air, using observed pressure, temperature, and relative humidity (RH). To calculate this ratio we use Tetens formula to determine the water vapor saturation pressure in the chamber (p_w, [kPa]) as a function of the observed air temperature (T_a, [C]),

$$p_w = 0.62198 * e^{\left(\frac{17.27*T_a}{T_a+273.3}\right)}, \quad (1)$$

and subsequently calculate the humidity ratio (x_s) given p_w and the atmospheric pressure measured empirically (p_a, [kPa]):

$$x_s = \frac{0.62198 * p_w}{p_a - p_w} \quad (2)$$

We then determine the humidity ratio for a given fractional relative humidity (RH, %, divided by 100) with RH measured at the sampling time:

$$x = x_s * (RH/100) \quad (3)$$

Using the humidity ratio and gas law constants, we convert observed $[CO_2]_{Obs}$ [ppm] to [kg/m^3]. The reference density of *dry* air [kg/m^3] is first calculated as

$$\rho_{ref} = \frac{p_a * 100}{R_a * (T_a + 273.15)} \quad (4)$$

Table 1
Parameters Needed for Calculating Flux

Parameter	Symbol	Units	Typical value
Chamber volume	V_c	m^3	0.002758
Chamber surface area	A_c	m^2	0.01455
Molar mass CO_2	M_{CO_2}	kg/mol	0.044009
Molar mass dry air	M_a	kg/mol	0.0289628
Specific gas constant, dry air	R_a	J kg^{-1} K^{-1}	287.058
Specific gas constant, water vapor	R_w	J kg^{-1} K^{-1}	462.5
Gas constant ratio, water to air	$\frac{R_w}{R_a}$	*Unitless*	1.609

where R_a is the gas constant of dry air (Table 1), and observed T_a [C] is converted to Kelvin. We use this reference density to calculate the moist air density [kg/m_3], using the humidity ratio calculated at this observation's temperature, pressure, and RH, and the gas constant ratio between water vapor and air (c.f. Table 1):

$$\rho_w = \rho_{ref} \frac{1+x}{1+\left[\frac{R_w}{R_a}\right] * x} \tag{5}$$

We calculate the molar concentration of moist air [mol/kg^3] according to

$$[air]_{mol} = \frac{1}{M_a} * \rho_w \tag{6}$$

where M_a is the molar mass of dry air (c.f. Table 1). Using the output of Equation 6, we convert CO_2 in [ppm] to molar density [mol CO_2/m^3 air] according to

$$[CO_2]_{mol} = air_{mol} * [CO_2]_{\text{Obs}} \tag{7}$$

and using the molar mass of CO_2 (c.f. Table 1) to convert molar density to mass density [kg CO_2/m^3 air], according to

$$\rho_{CO_2} = [CO_2]_{mol} * M_{CO_2} \tag{8}$$

Finally, the time varying mass of CO_2 in the chamber, $C(t)$ [kg], is found using the chamber volume, V_c [m^3] (c.f. Table 1), as

$$C(t) = \rho_{CO_2}(t) \times V_c, \tag{9}$$

Individual values of $\rho_{CO_2}(t)$, RH, and T_a are derived from 20-s averages of $[CO_2]_{\text{Obs}}$ (Text S6 and Figure S7 in Supporting Information S1). To eliminate the introduction of bias from seven fluxbots whose pressure sensors were compromised after a rainfall event in late August, for each hour's atmospheric pressure measurement we calculated the mean atmospheric pressure across the other five fluxbots in the array, and used these to calculate flux for each hour for all the fluxbots in the array (Text S6 in Supporting Information S1).

3.2. Key Parameters for the Fluxbot System

The parameters for the above flux calculations include values like "chamber volume" and "chamber surface area" that are specific to the design of these fluxbots and essential to calculation of flux for any chamber-based system. These also include constants related to the properties of CO_2, air, etc. (Table 1).

3.3. Flux Calculations

During an observation event, the system is sealed and the mass of CO_2 in chamber is monitored for 5 min. Throughout the 5-min observation, we track $C'(t)$ [kg], which is the time-varying mass of CO_2 relative to the initial mass, C_0 [kg], found according to

$$C'(t) = C(t) - C_0, \tag{10}$$

where C_0 is the initial mass of CO_2 in the chamber, which is derived from the ambient CO_2 concentration, $[CO_2]_A$ recorded just prior to the initiation of a measurement. We estimate the flow rate (mass change per unit time) of CO_2 into the chamber by fitting a first-order regression between every time of observation, t_i, and each ith observation of relative CO_2 mass, C'_i.

$$C'_i = \beta_0 + \beta_1 t_i + \epsilon_i \tag{11}$$

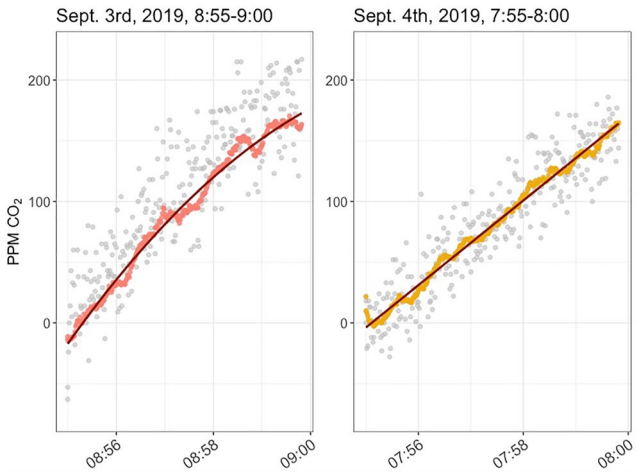

Figure 4. Two different measurement intervals taken by the same fluxbot (located in the "total wildlife exclusion" plot, at "open soil" site 2), 1 day apart at similar times in the morning. The left panel best fits a second-order polynomial regression, while the right panel best fits a linear regression. Gray points are raw data; salmon (left) or yellow (right) points are averaged with a 20 s rolling mean and the regression fitted to that averaged data. Data are transformed to begin at zero to emphasize the change in CO_2 concentration over time.

Because $C'_{i=0} \equiv 0$, it is given that the expected value of β_0, $E[\beta_0]$, is equal to 0 as long as we make the standard regression assumption that $E\left[\epsilon_i | C'_i\right] = 0$.

We estimate the average flow of CO_2 mass during each measurement period by taking the derivative of Equation 11 with respect to time, t, which yields

$$\frac{dC'}{dt} = \beta_1. \tag{12}$$

In many cases, CO_2 concentrations (and, therefore, values of C') in the chamber are changing in a non-linear manner over the course of each 5-min measurement interval. Because the CO_2 concentration in the chamber changes over the period of observation, we note that the observed flow rate of CO_2 into the chamber could also change as the gradient in $[CO_2]$ between the soil and the chamber shifts. This time-dependence in the evolution of $C'(t)$ is accounted for by using higher-order terms according our regression between C'_i and t_i.

$$C'_i = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \epsilon_i \tag{13}$$

$$= \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \beta_3 t_i^3 + \epsilon_i \tag{14}$$

As in the linear case (Equation 11), we can estimate the time-varying flow of CO_2 mass during each measurement period by taking the derivatives of Equations 13 and 14 with respect to t, which yield

$$\frac{dC'_i}{dt} = \beta_1 + 2\beta_2 t_i \tag{15}$$

$$= \beta_1 + 2\beta_2 t_i + 3\beta_3 t_i^2 \tag{16}$$

Because we are seeking to obtain the best possible estimate of the initial rate of CO_2 accumulation, we focus on our estimators for β_1 of both the first- and second-order regression, which will always describe the initial rate of change in C' within the chamber at the start of a measurement interval (i.e., when $t = 0$) (Hereafter, we will use "β_1" to refer to the initial rate of change for either first- or second-order regressions.). In the case where observed CO_2 accumulation rates are higher at the start of a measurement, the value of β_2 for a second-order regression will be less than zero. In contrast, when β_2 is greater than zero, it indicates an observation interval in which for some reason the rate of accumulation of CO_2 in the chamber is increasing during the measurement interval. Overall we find that of the 10,107 fluxes, only 3 were calculated with a slope from a second-order regression whose β_2 was *larger* than zero, indicating that for the overwhelming majority of fluxes the initial rate of change is the highest.

Equations 11, 15, and 16 provide three different examples of how to estimate β_1, the value of dC'/dt when $t = 0$. In designing our data processing software, for each observation interval we determine a β_1 estimate for *both* a first- and second-order polynomial regression. Between the two β_1 estimates, we select the larger with which to calculate flux for that interval. For closed-chamber systems, not all accumulation curves are best described by a linear regression (e.g., Figure 4); in cases where linear regression is inappropriately applied to non-linear accumulation curves, fluxes can subsequently be underestimated (Kutzbach et al., 2007). Finally, the flux of CO_2 into the chamber for each measurement interval, f_{CO_2} [kg m^{-2} s^{-1}] is determined as the estimated initial rate of change of CO_2 mass (the larger of the two β_1 [kg/s] values) divided by the surface area of the chamber, A_c [m^2] (c.f. Table 1):

$$f_{CO_2} = \frac{1}{A_c}\frac{dC'}{dt} = \frac{1}{A_c}\beta_1, \tag{17}$$

We conducted the above flux estimate calculations in Python, writing a custom script to batch-process our dataset's observation intervals (Forbes & Caylor, 2021). To calculate the regressions for each observation interval, we defined a regression fit function for which the user defines "n" as the order of the polynomial used to fit the raw data: one (hereafter linear regression), two (hereafter quadratic), or more. We normalized our CO_2 data to start at zero by subtracting the first CO_2 value from each subsequent value in the array. We therefore also defined our

function to fit regressions with an intercept of zero (Fitting regressions without an intercept can result in negative R^2 values when data are noisy.). We also defined an error function, which calculates the difference between each time point's observed CO_2 value and the predicted value from a regression.

Using the "optimize.leastsq" function from the "optimize" package of the SciPy library (with input parameters of the error function, an initial estimated value for the fit's slope, and the data to be fitted [i.e., t and C']), we identified the fit with the smallest sum of squares for an observation interval. Once the best regression fits for each interval were identified, we extracted and compared their initial slopes (i.e., β_1) and selected the larger for the flux calculation.

We defined a function that converted these β_1 estimates to flux rates using the parameters in Table 1, as described in Equation 17. We included functionality to convert from mass of CO_2 change over time and surface area ([kg m^{-2} s^{-1}]) to smaller metric units and also molar units (e.g., [μmol m^{-2} s^{-1}]). Using the estimates of β_1 and fitted values, we also calculated upper and lower 95% confidence intervals of each β_1 estimate (and thus maximum and minimum flux estimates), each regression's R^2 (or goodness-of-fit), and the relative uncertainty of each flux estimate, as described below.

3.4. Flux Confidence Intervals, Uncertainty

After determining β_1 estimates for each flux interval's regression, we calculated the sum of squared deviations (SS_{tt}) from the mean for the length of the interval according to

$$SS_{tt} = \sum_{i=1}^{n} \left(t_i - \bar{t}\right)^2 \tag{18}$$

with n being the interval length, in total number of seconds. We subsequently calculated the variance of the regression, or the predicted values of CO_2 concentration at each time point in the flux interval compared to the observed values, according to

$$s^2_{C',t} = \left(\frac{1}{n-2}\right) \sum_{i=1}^{n} \left(C'_i - \hat{C}'_i\right)^2 = \frac{SS_E}{n-p} \tag{19}$$

where $n - p$ is the least square's degrees of freedom (with p the number of parameters used to calculate the regression; 2 or 3, for linear and quadratic regressions respectively) and SS_E, or the error sum of squares, is calculated with

$$SS_E = \sum_{i=1}^{n} \left(f - C'_i\right)^2 \tag{20}$$

when f is the predicted value of CO_2 at a given time step i. Using $s^2_{C',t}$ we then calculated the variance of β, according to

$$s^2_{\beta} = \frac{s^2_{C',t}}{SS_{tt}} \tag{21}$$

where the numerator is determined by Equation 19 and the denominator by Equation 18.

We subsequently determined the upper and lower 95% confidence intervals for each regression's β_1 estimate according to

$$\beta_1 \text{ConfidenceInterval}(\pm 95\%) \cong \beta_1 \pm 2\sqrt{s^2_{\beta_1}} \tag{22}$$

Using the upper and lower confidence intervals and the initial β_1 estimate we could calculate a minimum and maximum flux for each flux estimate. With these values we were able to determine the relative flux uncertainty of each flux estimate, first by calculating the absolute flux uncertainty according to

$$\text{absoluteuncertainty} = \left|\frac{\text{maximumfluxestimate} - \text{minimumfluxestimate}}{2}\right| \tag{23}$$

Table 2
The Total Number of Removed Flux Estimates After Quality Assurance/Quality Control (725), Broken Down by the Reasons for Which They Were Removed (e.g., Which of the Seven Errors Were Flagged), the Total Accumulated Error Score Associated With Those Reasons, and the Number Removed for Each

Reason for removal	QAQC flags	Number removed
3	100	10
4	1,000	16
5 and 3	10,100	429
5 and 3 and 2	10,110	11
5 and 4	11,000	1
6	100,000	248
6 and 2	100,010	7
6 and 4	101,000	3

We divided this absolute uncertainty by the flux estimate itself; this multiplied by 100 is the uncertainty of the flux estimate relative to its size, or its relative uncertainty. To account for negative flux estimates, we took the absolute value of relative flux uncertainty.

We calculated R^2 of each regression for all observation intervals. We first calculated SS_{yy}, the total sum of squares of the difference between the observed values of CO_2 concentration and the average value of CO_2, according to

$$SS_{C'C'} = \sum_{i=1}^{n} \left(C_i' - \hat{C}'\right)^2 \quad (24)$$

We used the output as a term in the numerator and denominator for calculating each regression's R^2, and the error sum of squares from Equation 20, according to

$$R^2 = \frac{SS_{C'C'} - SS_E}{SS_{C'C'}} \quad (25)$$

4. Quality Assurance/Quality Control (QA/QC)

We removed 1,366 flux observations with incorrect timestamps. In addition, we treated the three weeks of data collection that occurred between 2 August 2019 (fluxbot installation) and 23 August 2019 (on which date the fluxbots were re-calibrated) as a test period, and removed these data from the final dataset. This period occurred immediately after the fluxbot array's installation and therefore likely included time during which the soils were disturbed (and flux could be artificially high for days or weeks [Davidson et al., 2002]). Additionally, this period included post-installation troubleshooting that involved some disturbance, like several unusually early rainfall events in August that required we remove the electronics systems from each fluxbot and paint the connections with clear nail polish as a conformal coating to improve weatherproofing.

We removed all the data collected in October for a single fluxbot whose hardware shorted at the end of September 2019, likely due to water damage from one of several heavy rain storms that occurred around that time. We continued collecting data at that site by rotating lids from "donor" fluxbots (e.g., a randomly-selected fluxbot's lid was removed and placed on the broken fluxbot's chamber) for 2–3 days each. These 229 fluxes were analyzed separately to explore fluxbot variability at the same site over time.

Finally, to identify and remove any remaining spurious flux calculations (e.g., calculated with compromised raw data), we implemented an automated quality assurance and quality control (QA/QC) process in the data processing software. All flux intervals were examined for a total of seven possible errors, listed in order of severity:

1. *net change in atmospheric pressure greater than* 10 hPa;
2. *net change in temperature greater than 2.*5C;
3. *extraordinary maximum* CO_2 *values (e.g., that could be expected from errors ranging from electronic failure to heterotrophic respiration inside the chamber from a trapped invertebrate)*;
4. *whether the measurement interval encompassed a minimum of 4.5 min, or 270 total observations*;
5. *if the net change in* CO_2 *over a measurement interval was less than* 10 ppm*, indicating possible leakage or imperfect seal*;
6. *if the net difference in* CO_2 *from start to finish of a measurement interval was negative, indicating a greater possible likelihood of leakage or* CO_2 *uptake from an errant photosynthesizing plant inside the chamber*;
7. *if the last recorded value of* CO_2 *was less than the mean, and the first recorded value was greater than the mean.*

These seven errors were each assigned a value of [1 * an increasing order of magnitude]; a single flux estimate could therefore "earn" an error score of up to 1,111,111 (e.g., adding 1, 10, 100, 1,000, 10,000, 100,000, and 1,000,000). Any flux estimate with an error score of more than 11 was discarded (e.g., fluxes that had either high pressure *or* temperature buildup, as well as those that had *both* high pressure *and* temperature buildup, were kept in the final dataset). We removed a total of 725 flux estimates for failing QA/QC (Table 2; Figure S8 in Supporting Information S1).

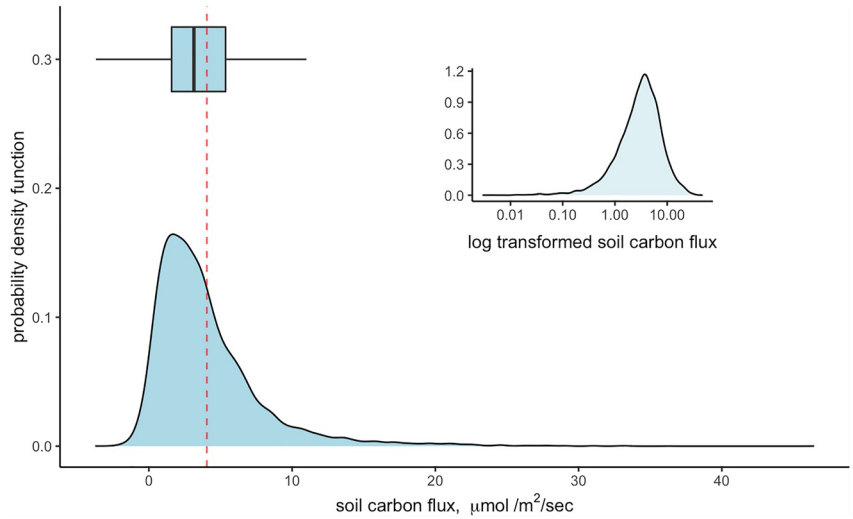

Figure 5. Density distribution of fluxes from entire deployment period. The inset boxplot represents the dataset's dispersion and skew, with the central vertical line indicating the median (3.14 µmol/m^2/s); all points in the distribution located to the right of the maximum observed non-outlier value (the right "whisker") are, statistically, abnormally high occurrences of flux. The inset lognormal density distribution shows the distribution of the (positive only) log-transformed fluxes. The vertical red dashed line indicates the dataset's overal mean (4.05 µmol/m^2/s).

5. Results

Our array of 12 fluxbots collected 10,107 quality-controlled estimates of soil CO_2 flux. Mean flux across all 12 fluxbots, over time, was 4.05 µmol/m^2/s, and ranged from −3.70 to 46.40 µmol/m^2/s. Of the 10,107 flux estimates, 334 were below zero; however of those negative fluxes, the vast majority (87%) occurred between zero and −1, aka clustered around zero (Figure 5). Two-thirds of the fluxes were estimated with the quadratic regression's β_1; these fluxes were generally higher than those calculated using the linear regression's β_1 (Figure 6).

5.1. Patterns of Flux

Raw CO_2 concentration data demonstrated distinct diel patterns over time (Figure 7a), with atmospheric CO_2 higher in nighttime hours than day. Fluxes *also* demonstrated distinct diel patterns, with the highest fluxes occurring at night and lowest at midday or mid-afternoon (Figure 7b); a simple two-sided *t*-test comparing daytime fluxes (8 a.m.

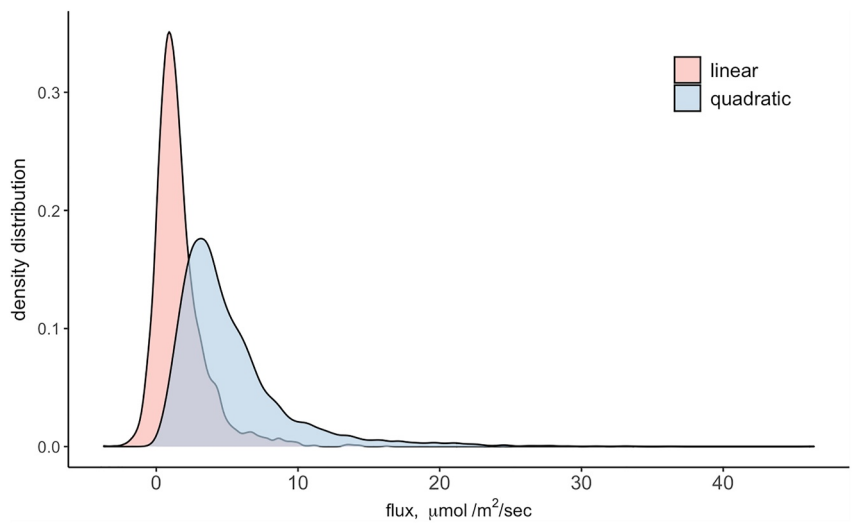

Figure 6. Density distribution of fluxes calculated with a β_1 value derived from a linear regression versus a quadratic regression.

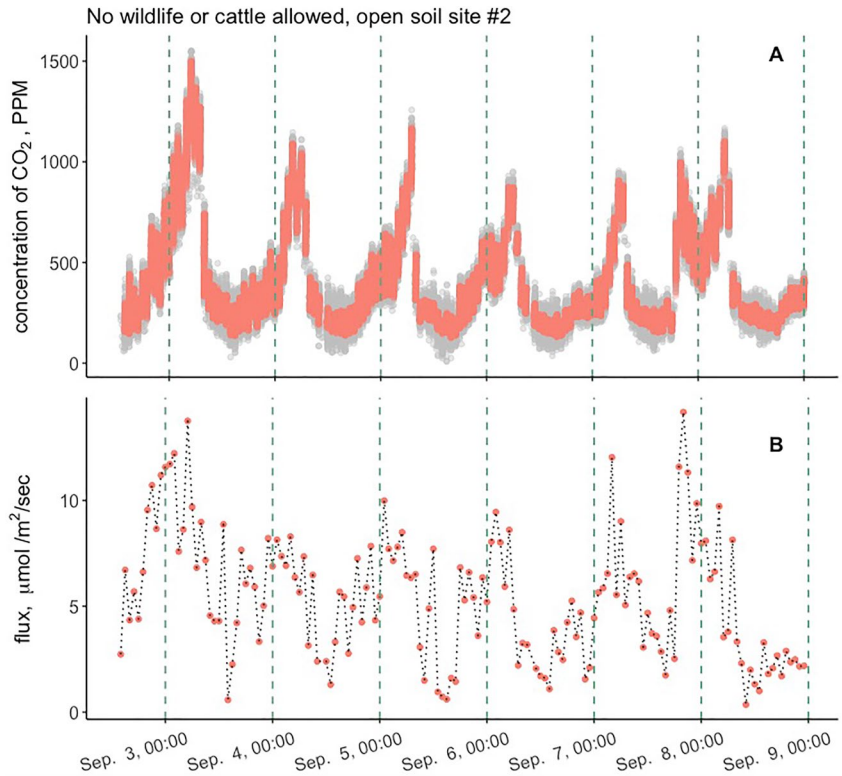

Figure 7. (a) Raw CO_2 concentration data, collected from a single fluxbot over 6 days in September 2019. Gray points are untransformed raw CO_2; salmon points are transformed with the 20 s averaging window. (b) Fluxes, calculated from the 20 s averaged CO_2 data. For both panels, green dotted lines indicate each day's midnight hour.

to 4 p.m., or 2 hr after sunrise and 2 hr before sunset) and nighttime fluxes (8 p.m. to 4 a.m.) revealed consistently higher fluxes in nighttime hours (mean 4.41 µmol/m^2/s) than in daytime hours (mean 3.74 µmol/m^2/s) ($p < 0.001$).

Fluxes collected from beneath *A. drepanolobium* canopies (mean 3.69 µmol/m^2/s) were lower on average than those collected from open soil patches (mean 4.31 µmol/m^2/s) ($p < 0.001$); this difference was driven by time of day, with nighttime fluxes increasing significantly at open soil patches (Figure S9 in Supporting Information S1). Fluxes collected from the treatment plot open to all large-bodied herbivores (mean 4.843 µmol/m^2/s) were consistently higher than those collected from the fenced exclosure plot (mean 3.589 µmol/m^2/s) ($p = 0.001$).

5.2. The "Rotating Lid" Experiment

As described in Section 4, we removed 229 fluxes from the dataset due to their being collected at a fluxbot that received "donor" lids every 2–3 days for the month of October. We visualized these data separately to explore variability in fluxbot lid performance at the same location. Ambient CO_2 detection varied by donor lid, indicating between-fluxbot variation in absolute CO_2 detection across the array. However, fluxes were consistently in the same range during the entire period, no matter the identity of the donor lid and their ambient CO_2 detection accuracy (Figure S12 in Supporting Information S1).

5.3. Relative Flux Uncertainty and Model Fits

We visualized each flux estimate against its relative flux uncertainty, and its regression's R^2 value. We found that relative flux uncertainty was generally low, and that the highest uncertainties tended to be associated with fluxes at, near, or below zero (Figure S11a in Supporting Information S1). Indeed, the vast majority (92%) of our flux estimates had relative flux uncertainties lower than 25%. While 164 flux uncertainties were extremely high (over 100%), 58 of their associated flux estimates were negative, and 151 were associated with fluxes less than 0.15 µmol/m^2/s (i.e., clustered extremely close to zero).

Seventy-one percent of all flux estimates were associated with an R^2 value greater than 0.90, indicating generally high "goodness of fit" of the regressions selected to estimate flux (Figure S11c in Supporting Information S1). 9% of flux estimates were associated with negative R^2 values. However, over one third of the flux estimates associated with these negative R^2 values clustered around zero (e.g., between 0 and −1 μmol/m^2/s). A Spearman's correlation between R^2 and flux estimate was relatively high (correlation coefficient = 0.62). The mean R^2 associated with fluxes calculated using a linear β_1 was 0.47, while mean R^2 associated with those calculated using a quadratic β_1 was 0.73, indicating that quadratic regressions were generally better fits for the data of steeper CO_2 accumulation curves (also see Figure 6).

Eighty-six percent of flux estimates associated with lower (here, less than 0.5) R^2 values nonetheless also had low relative flux uncertainty (again, less than 0.5) (Figure S11b in Supporting Information S1). Only 3.4% of all flux estimates had *both* a low R^2 and high relative uncertainty.

5.4. Comparison With Manual Data Collection

We were unable to directly compare the data we collected using the manual CIRAS flux chamber to the fluxbot dataset, given that the measurements were not taken on the same footprint of soil. Nevertheless we conducted a comparison of the two datasets to explore mechanisms of difference. The mean for CIRAS observations paired in time (within 25 min) and space (within 20 cm) with fluxbot observations was 1.58 μmol/m^2/s, with a median of 1.38 μmol/m^2/s; the coefficient of variation was 0.47. The mean for fluxbot-collected fluxes matched with those collected manually was 3.50 μmol/m^2/s, with a median of 2.56 μmol/m^2/s; the coefficient of variation was 0.98 (i.e., a larger spread) (Figure S13 in Supporting Information S1). A Goodman-Kruskal gamma correlation, which reduces the influence of outliers (e.g., "hotspots" of flux), indicated a weakly negative correlation between the two methods (−0.20, $p = 0.001$).

6. Discussion

We collected 10,107 flux observations over approximately two months with the fluxbot array. To our knowledge, this is the largest and most comprehensive dataset of soil carbon flux of this ecosystem. In KLEE, it takes an average of 10 min to collect an observation using a manually-operated commercial survey chamber, including travel between collars (Forbes, personal observation). It would therefore take over 1,680 hr to manually collect the same number of observations, which would still exclude key time periods during which manual collection is risky or infeasible (night, after rainfall). Previous studies here therefore prioritized either high spatial resolution and extent (Forbes unpublished data), or repeated measures at fewer sites (DeCarlo & Caylor, 2020).

We were also able to distribute our array in a long-running community ecology experiment, and across distinct ecological structural features known to be mediated by the presence or absence of large wildlife (e.g., Charles et al., 2017; Young et al., 2018). This strategy demonstrates that a distributed array can be leveraged by community and wildlife ecologists, allowing the broader ecology community to "animate the carbon cycle" (Schmitz et al., 2018) by interrogating datasets collected across community-level experimental contexts, climate or rainfall gradients, land-use change or wildlife loss, space-for-time successional gradients, and more.

We observed significantly higher average flux rates at night than during the day (Figure 7b). While unusual, such a diel pattern is not unheard of in tropical ecosystems, where soil carbon flux may be more tightly coupled to moisture than to temperature (Otieno et al., 2011; Zhao et al., 2021). Given that dew is a significant source of moisture in this savanna (Ngatia et al., 2014), evening dew formation could be acting as a daily "drying and re-wetting" mechanism, spurring higher fluxes at night (Figure S10 in Supporting Information S1). It is also likely that the characteristic surface and subterranean cracking of this vertisol mechanically facilitates thermal convection when ambient air temperatures drop (Figure S16 in Supporting Information S1), further enhancing nighttime fluxes and perhaps producing dynamic "hotspots" of high flux (DeCarlo & Caylor, 2020).

Within this cyclical diel pattern, we observed that deployment location influenced flux rates. Fluxbots located in open areas detected higher fluxes than those underneath *A. drepanolobium* trees, and particularly at night (Figure S9 in Supporting Information S1). Fluxbots in the unfenced plot, open to large-bodied herbivores, detected higher fluxes than those in the fenced exclosure plot. While not in the scope of this paper, we can theorize on the interacting biological, geological, and edaphic mechanisms driving these patterns. For example, biocompaction from

megafauna like elephants produces deep, layered vertical cracks in the soil, thus higher average fluxes and more frequent high outliers (DeCarlo & Caylor, 2020). We speculate that since biocompaction from animal movement in the unfenced plot is greater than in the fenced herbivore exclusion plot, and is also likely greater in open areas than it is under tree canopies, it facilitates higher average and maximum fluxes and amplifies nighttime thermal convection effects in those locations.

6.1. Comparison With Manual Survey Data Collection

Due to differences in the shape and size of the fluxbot and CIRAS chambers, it was not possible to validate fluxbot performance by manually measuring flux at each fluxbot's exact footprint (Savage & Davidson, 2003) and we relied on measurements taken adjacent to each fluxbot for comparison (Text S7 in Supporting Information S1). These measurements, though adjacent in space and time, were ultimately not comparable (Figure S13 in Supporting Information S1). Different shapes and sizes of chambers can also result in different estimates of the same flux (Pumpanen et al., 2004), making direct comparison even on the same footprint difficult.

However, high spatial variability in soil carbon flux is not uncommon (Fóti et al., 2016; Wang et al., 2021). Markedly different, simultaneous fluxes can occur centimeters apart (Davidson et al., 2002), with sub-centimeter "hot spots" due to dynamic incursions: soil aggregates, gaps or cracks, and soil fauna activity or engineering (Ohashi et al., 2007; Stoyan et al., 2000). The clay-rich vertisol in KLEE is particularly dynamic, forming cracks from drying (Somasundaram et al., 2018) and faunal activity (DeCarlo & Caylor, 2019), which can locally increase soil carbon flux by orders of magnitude (DeCarlo & Caylor, 2020) even between spatially close (less than 15 cm) sampling sites (Rochette et al., 1991). Given the characteristics of KLEE soils we suggest that "hot spot"-inducing factors likely account for some lack of fidelity between fluxbot and CIRAS estimates.

An alternative hypothesis is that the array included "hot bots" whose sensors erroneously detected steeper gains in raw CO_2 than the higher-accuracy sensor in the CIRAS. To determine whether fluxbots that consistently detected large daily ranges in atmospheric CO_2 concentration *also* consistently estimated higher fluxes than the CIRAS, we plotted the fluxbots' average daily CO_2 amplitude (to characterize each fluxbot's sensitivity to change in CO_2) against the ratio of its paired fluxbot and CIRAS flux estimates. However, we did not see such evidence of "hot bots" (Figure S14 in Supporting Information S1). Those fluxbots with the largest daily ranges were associated with flux estimates that were comparable to, or even underestimated, CIRAS estimates. Some fluxbots with lower daily ranges were also associated with relatively comparable flux estimates, but others were associated with fluxbot estimates that were (on average) larger than their paired CIRAS estimates. We deduced that whatever is causing dissonance between fluxbot and CIRAS estimates is likely biological or environmental variability (e.g., "hot spots"), as described above, than "hot bots" with over- (or under-) sensitive CO_2 sensors.

6.2. Uncertainty and Percent Error in Flux

We reason that *initial* rates of increase are the least likely to result in an underestimate of true flux, due to the possibility of pressure buildup inside a closed chamber reducing the diffusion gradient of CO_2 from the soil over time (Baneschi et al., 2023; Kutzbach et al., 2007; Pumpanen et al., 2004). We thus assumed that the higher initial rate of increase (β_1) between the linear and quadratic regressions of a given CO_2 accumulation interval is most representative of real flux, regardless of the chosen regression's R^2 (but see Carbone et al., 2008, 2011). Flux estimates whose associated R^2 is low may simply indicate a noisy accumulation interval, resulting from environmental variability or other unobserved variables. Alternatively, a low R^2 could indicate that a regression's slope was close to or at zero, when low R^2 values are statistically certain. This includes the few (9%) negative R^2 values in the final dataset; because we normalized raw CO_2 accumulation intervals to begin at zero, we also set all regressions' intercepts to zero. With noisy data, for which subtracting the initial value of CO_2 may (e.g.) have resulted in data scattered *around* zero, the resulting regression line may be forced away from a fit with the lowest possible error sum of squares.

We argue that relative flux uncertainty, predicated as it is on *initial* rates of CO_2 accumulation rather than a regression's goodness of fit, provides a more meaningful assessment of flux estimate accuracy. 86% of our flux estimates associated with low R^2 values had fairly low relative flux uncertainty (under 0.5). However, estimates with high relative uncertainty are still likely to be representative, but of low flux rates; CozIR LP Miniature sensors have a typical accuracy of $\pm$30 ppm, meaning that shallow accumulation curves will produce worse-fitting regressions and higher relative error. As such we also did not eliminate flux estimates with high relative flux uncertainty, considering that the few flux estimates with high relative uncertainty were close to zero. While very

low flux estimates may therefore have more uncertainty, we contend that they are no less representative and that removing them would unduly bias the dataset toward higher fluxes.

6.3. Inter-Fluxbot Variability in Ambient CO_2 Detection

There is clearly variability across the sensors in their ability to detect *absolute* CO_2, and sensor drift over time (Text S8 and Figure S15 in Supporting Information S1). However, the distribution of *flux estimates* collected with the fluxbots (Figure 5) is consistent with those from studies of this system and other tropical sites (Courtois et al., 2019; DeCarlo & Caylor, 2020; February et al., 2020; Konaté et al., 2003; Poth et al., 1995). In addition, in our study of rotating "donor" lids at a single chamber location in October 2019 (Text S8 in Supporting Information S1), randomly-selected lids detected consistent fluxes despite also collecting characteristically different raw CO_2 concentration data (Figure S12 in Supporting Information S1). As fluxes are calculated by integrating *net* change in CO_2 over time, this experiment demonstrated that the fluxbots are consistent in observing net change despite variation in absolute CO_2 detection (Bastviken et al., 2015).

Also using a CozIR brand sensor, Helm et al. (2021) observed offsets in individual sensors' measurements of raw CO_2 concentrations in the field. However, they also observed that net changes in CO_2 were consistently observed between sensors with differing ambient baselines (Helm et al., 2021). While we did not test each fluxbot lid on each fluxbot body, given the results of our donor lid experiment and support from the literature we feel confident in the fluxbots' ability to accurately capture soil carbon flux.

6.4. Further Development

The goal of our fluxbot array was to improve data collection capacities across large spatial extents with high spatial and temporal heterogeneity. While we have demonstrated the fluxbot array's utility in achieving that goal, there are key faults to our design that can be improved upon. For example, the CO_2 sensors demonstrated significant offsets in ambient CO_2 detection across the array. While our "hot bot" analysis demonstrated that it was not likely in this case (Figure S14 in Supporting Information S1), variation in offset within an array could result in variable detection of the same dCO_2, given the possibility for nonlinear absorbance of CO_2 in different models of inexpensive NDIR sensor (see Yasuda et al., 2012).

We suggest that, no matter the sensor chosen, adopters of our design adjust the software's collection of ambient CO_2 concentration to one single period of at least 20 s (per our Allan variance analysis) prior to chamber closure. These improved ambient atmospheric CO_2 data could be used to cross-validate ambient CO_2 detection across every fluxbot deployed in a landscape, leveraging the possibilities of a networked array yoked by a communication hub that can receive data from and send updates to each fluxbot. To further harden the system against moisture or inclement weather, we suggest conformal coating, dip-coating, or shrink-wrapping the finished electronics systems' connections, and using a more weather-resistant temperature and humidity sensor (e.g., Adafruit's SHT-30 with a tightly-woven mesh cover).

We calibrated our sensors with a single-point calibration and using the CIRAS and SRC-2 chamber to ensure that we could recalibrate with the identical method in the field, where an airtight chamber and canisters of CO_2 standards are difficult to procure and transport. Due to the observed variability in absolute CO_2 detection across the array, however, we recommend that adopters of our design conduct an initial in-lab calibration of each sensor with known standards. We also recommend conducting regular recalibrations every two to 3 weeks (Helm et al., 2021). Adopters of our design could also use the "auto-zero" setting available on CozIR brand sensors, which when enabled, re-zeros the sensor daily at a predetermined time during which atmospheric CO_2 concentration is stable; we suggest performing field reconnaissance prior to fluxbot deployment to identify an appropriate re-zero time.

To monitor sensor drift at a slightly higher overall cost, we suggest that a random subset of fluxbots in an array receive a duplicate CO_2 sensor, against which the main sensor could be compared and monitored for drift. Lastly, while we did not observe noticeable drift in time across our array, to ensure continued temporal comparability we suggest that researchers collecting data over longer time periods determine whether RTC drift occurs before deployment.

We experienced significant difficulty in producing a dataset of CIRAS-derived flux estimates against which to validate fluxbot-derived flux estimates. As a result, these datasets were not directly comparable. We suggest that

researchers seeking to validate their fluxbots' performance consider an adapter collar, fitted on one side to the size and shape of the commercial chamber and on the other to the fluxbot body, to take data on the same footprint. Alternatively (and given the inherent difficulties in comparing flux estimates taken with chambers of different provenance [Pumpanen et al., 2004]) we recommend the use of more adaptable gas exchange instruments (e.g., LI-COR 7810 or Los Gatos Research systems), and connecting tubing from the instrument itself to the fluxbots' chamber, to take measurements directly from the fluxbots' own internal volume.

6.5. Benefits of the Fluxbot Array

CozIR and other inexpensive sensors have been used extensively in instrument development and data collection from greenhouses to forests (Changqing et al., 2018; Folea & Mois, 2015; Gibson & MacGregor, 2013; Gong et al., 2013; Helm et al., 2021; Wen & Li, 2013). The drawbacks to inexpensive sensors are well known, and can be adjusted for with more widespread field-based experimentation and improvements like those described above. We argue that the advantages of inexpensive sensors for distributed soil carbon flux sensing vastly outweigh the costs.

Soil carbon flux datasets that capture small-scale heterogeneity can also validate and reduce uncertainty as parameters in large-scale models of carbon cycling (Keenan et al., 2012). While we designed this tool from the perspective of community ecologists, whose goals increasingly include connecting biodiverse animal communities to carbon cycling, we also argue that the approach can be useful to scientists seeking to better parameterize larger-scale carbon models.

An array of independently-operating fluxbots can be expansive at a fraction of the cost of commercial chambers, increasing researchers' capacity for high-resolution data collection and their statistical likelihood of converging on a landscape's mean representative flux (despite the possibility of increased variability for a single flux observation). Critically, this approach offers the opportunity for ecologically-relevant, systematic deployment strategies that capture small-scale heterogeneity within functionally complex environments. Expanding the scales at which we collect carbon flux datasets could illuminate the mechanisms by which communities respond to and produce heterogeneity in carbon dynamics (Chave, 2013), and be another method for "animating the carbon cycle" (Schmitz et al., 2018).

6.6. Conclusions

Using the fluxbot array, we identified an unusual pattern of high nighttime flux driven by edaphic characteristics of this savanna. Deploying fluxbots at high spatial resolution, at ecologically-distinct landscape features, and within a longstanding animal community experiment allowed us to further consider the mechanisms that drive fine-scale variability in soil carbon flux, including direct (compaction and bioturbation [DeCarlo & Caylor, 2019]) and indirect (changes to ecosystem structure like tree abundance [Sitters et al., 2020]) effects of the wildlife community. Here we demonstrate a path forward for inexpensive collection of ecologically "big" soil carbon flux data.

The fluxbot design and processing software is openly accessible. We hope that other researchers' replication of our approach will demonstrate that high-resolution data collection is accessible even on a limited project budget. The generation of highly-resolved datasets with distributed fluxbot arrays, particularly in understudied ecosystems, will better our collective understanding of how patterns of soil carbon flux at small scales contribute to ecosystem carbon dynamics globally (Chave, 2013).

Data Availability Statement

The raw data generated by the fluxbots, as well as the processed flux data calculated from these raw data, is available at the archived project Github repository (repository found at the Zenodo DOI: https://doi.org/10.5281/zenodo.7650956) (Forbes & Caylor, 2021). Software that runs the fluxbots can be found at the project repository in a submodule run by coauthor VB, labeled "hardware." All figures were generated in R, scripts for which can be found at the repository, with an R project (.Rproj) in a folder labeled "figures_Rscripts." Software to process flux data from raw data inputs can also be found at the repository in the main project ("fluxbot.py"). Also in the main project are scripts used to calculate Allan variance, generate the fluxbot events table from raw data, and compare fluxbot and CIRAS data; a PDF of the PP Systems CIRAS manual; and a README.md.

Acknowledgments
Field work was conducted at Mpala Research Centre and Conservancy, with the Republic of Kenya's National Commission for Science, Technology, and Innovation research permit numbers NACOSTI/P/16/18316/10582 and NACOSTI/P/19/523. We thank the Mpala Research Centre for hosting the research team in the field, and the Mpala staff for essential logistical support without which this work would not be possible. We thank Drs. Truman Young and Duncan Kimuyu for use of KLEE for this project. KLEE plots were built and maintained by grants from the Smithsonian Institution, The National Geographic Society (Grants 4691-91 and 9106-12), the African Elephant Program of the US Fish and Wildlife Service (98210-0-G563) and the National Science Foundation (LTREB BSR-97-07477, 03-16402, 08-16453, 12-56004 and 12-56034; DEB 19-31224). Research funds for supporting GL, VB, and SF were provided by an NSF Research Experience for Undergraduates award (NSF-REU 1820728) and an NSF Water Sustainability and Climate grant (NSF-WSC 1801251). Funds for the beta version of the fluxbot array were awarded to EF by the National Science Foundation Integrative Education and Research Traineeship (NSF-IGERT) at UC Santa Barbara. Research funds for supporting travel to and residence at Mpala, and salary support for research associates, was awarded to EF by a 2019 Schmidt Family Foundation Research Accelerator Award. We wholeheartedly thank Alfred Ekaaz for his invaluable contributions in the field testing the beta version, and for his essential experience with electronics in the field.

References

Adachi, M., Bekku, Y. S., Konuma, A., Kadir, W. R., Okuda, T., & Koizumi, H. (2005). Required sample size for estimating soil respiration rates in large areas of two tropical forests and of two types of plantation in Malaysia. *Forest Ecology and Management*, *210*(1), 455–459. https://doi.org/10.1016/j.foreco.2005.02.011

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*(11), 2415–2434. https://doi.org/10.1175/1520-0477(2001)082⟨2415:FANTTS⟩2.3.CO;2

Baneschi, I., Raco, B., Magnani, M., Giamberini, M., Lelli, M., Mosca, P., et al. (2023). Non-steady-state closed dynamic chamber to measure soil CO_2 respiration: A protocol to reduce uncertainty. *Frontiers in Environmental Science*, *10*, 1048948. https://doi.org/10.3389/fenvs.2022.1048948

Bastviken, D., Sundgren, I., Natchimuthu, S., Reyier, H., & Gålfalk, M. (2015). Technical Note: Cost-efficient approaches to measure carbon dioxide (CO_2) fluxes and concentrations in terrestrial and aquatic environments using mini loggers. *Biogeosciences*, *12*(12), 3849–3859. https://doi.org/10.5194/bg-12-3849-2015

Carbone, M. S., Still, C. J., Ambrose, A. R., Dawson, T. E., Williams, A. P., Boot, C. M., et al. (2011). Seasonal and episodic moisture controls on plant and microbial contributions to soil respiration. *Oecologia*, *167*(1), 265–278. https://doi.org/10.1007/s00442-011-1975-3

Carbone, M. S., Winston, G. C., & Trumbore, S. E. (2008). Soil respiration in perennial grass and shrub ecosystems: Linking environmental controls with plant and microbial sources on seasonal and diel timescales. *Journal of Geophysical Research*, *113*, g02022. https://doi.org/10.1029/2007JG000611

Changqing, C., Hui, L., & Wenjun, Y. (2018). Design and development of a greenhouse remote monitoring system based on wince. *Meteorological and Environmental Research*, *9*(4), 106–111. https://doi.org/10.19547/j.issn2152-3940.2018.04.024

Charles, G. K., Porensky, L. M., Riginos, C., Veblen, K. E., & Young, T. P. (2017). Herbivore effects on productivity vary by guild: Cattle increase mean productivity while wildlife reduce variability. *Ecological Applications*, *27*(1), 143–155. https://doi.org/10.1002/eap.1422

Chave, J. (2013). The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters*, *16*(s1), 4–16. https://doi.org/10.1111/ele.12048

Courtois, E. A., Stahl, C., Burban, B., Van den Berge, J., Berveiller, D., Bréchet, L., et al. (2019). Automatic high-frequency measurements of full soil greenhouse gas fluxes in a tropical forest. *Biogeosciences*, *16*(3), 785–796. https://doi.org/10.5194/bg-16-785-2019

Davidson, E. A., Richardson, A. D., Savage, K. E., & Hollinger, D. Y. (2006). A distinct seasonal pattern of the ratio of soil respiration to total ecosystem respiration in a spruce-dominated forest. *Global Change Biology*, *12*(2), 230–239. https://doi.org/10.1111/j.1365-2486.2005.01062.x

Davidson, E. A., Savage, K., Verchot, L. V., & Navarro, R. (2002). Minimizing artifacts and biases in chamber-based measurements of soil respiration. *Agricultural and Forest Meteorology*, *113*(1–4), 21–37. https://doi.org/10.1016/S0168-1923(02)00100-4

DeCarlo, K. F., & Caylor, K. K. (2019). Biophysical effects on soil crack morphology in a faunally active dryland vertisol. *Geoderma*, *334*, 134–145. https://doi.org/10.1016/j.geoderma.2018.07.042

DeCarlo, K. F., & Caylor, K. K. (2020). Effects of crack morphology on soil carbon flux dynamics in a dryland vertisol. *Geoderma*, *375*, 114478. https://doi.org/10.1016/j.geoderma.2020.114478

Delgado-Baquerizo, M., Eldridge, D. J., Maestre, F. T., Karunaratne, S. B., Trivedi, P., Reich, P. B., & Singh, B. K. (2017). Climate legacies drive global soil carbon stocks in terrestrial ecosystems. *Science Advances*, *3*(4), e1602008. https://doi.org/10.1126/sciadv.1602008

February, E., Pausch, J., & Higgins, S. I. (2020). Major contribution of grass roots to soil carbon pools and CO_2 fluxes in a mesic savanna. *Plant and Soil*, *454*(1), 207–215. https://doi.org/10.1007/s11104-020-04649-3

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Folea, S. C., & Mois, G. (2015). A low-power wireless sensor for online ambient monitoring. *IEEE Sensors Journal*, *15*(2), 742–749. https://doi.org/10.1109/JSEN.2014.2351420

Forbes, E., & Caylor, K. (2021). Fluxbots: A method for building, deploying, collecting and analyzing data from a network of inexpensive, autonomous soil carbon flux chambers. *Zenodo*. Retrieved from https://zenodo.org/badge/latestdoi/439441210

Fóti, S., Bálogh, J., Herbst, M., Papp, M., Koncz, P., Bartha, S., et al. (2016). Meta-analysis of field scale spatial variability of grassland soil CO_2 efflux: Interaction of biotic and abiotic drivers. *CATENA*, *143*, 78–89. https://doi.org/10.1016/j.catena.2016.03.034

Fox-Dobbs, K., Doak, D. F., Brody, A. K., & Palmer, T. M. (2010). Termites create spatial structure and govern ecosystem function by affecting N_2 fixation in an East African savanna. *Ecology*, *91*(5), 1296–1307. https://doi.org/10.1890/09-0653.1

Gibson, D., & MacGregor, C. (2013). A novel solid state non-dispersive infrared CO_2 gas sensor compatible with wireless and portable deployment. *Sensors*, *13*(6), 7079–7103. https://doi.org/10.3390/s130607079

Gong, H., Yu, H. L., Chen, G. F., & Wen, Z. (2013). *Design of measurement and control system of facilities vegetables based on internet of things* (Vol. 263–266). Trans Tech Publications Ltd. https://doi.org/10.4028/www.scientific.net/AMM.263-266.2824

Harmon, T. C., Dierick, D., Trahan, N., Allen, M. F., Rundel, P. W., Oberbauer, S. F., et al. (2015). Low-cost soil CO_2 efflux and point concentration sensing systems for terrestrial ecology applications. *Methods in Ecology and Evolution*, *6*(11), 1358–1362. https://doi.org/10.1111/2041-210X.12426

Helm, J., Hartmann, H., Göbel, M., Hilman, B., Herrera Ramírez, D., & Muhr, J. (2021). Low-cost chamber design for simultaneous CO_2 and O_2 flux measurements between tree stems and the atmosphere. *Tree Physiology*, *41*(9), 1767–1780. https://doi.org/10.1093/treephys/tpab022

Holden, J. (2005). Peatland hydrology and carbon release: Why small-scale process matters. *Philosophical Transactions of the Royal Society A*, *363*(1837), 2891–2913. https://doi.org/10.1098/rsta.2005.1671

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., et al. (2021). A restructured and updated global soil respiration database (SRDB-V5). *Earth System Science Data*, *13*(2), 255–267. https://doi.org/10.5194/essd-13-255-2021

Keenan, T. F., Davidson, E., Moffat, A. M., Munger, W., & Richardson, A. D. (2012). Using model-data fusion to interpret past trends, and quantify uncertainties in future projections, of terrestrial ecosystem carbon cycling. *Global Change Biology*, *18*(8), 2555–2569. https://doi.org/10.1111/j.1365-2486.2012.02684.x

Konaté, S., Le Roux, X., Verdier, B., & Lepage, M. (2003). Effect of underground fungus-growing termites on carbon dioxide emission at the point- and landscape-scales in an African savanna. *Functional Ecology*, *17*(3), 305–314. https://doi.org/10.1046/j.1365-2435.2003.00727.x

Kutzbach, L., Schneider, J., Sachs, T., Giebels, M., Nykänen, H., Shurpali, N. J., et al. (2007). CO_2 flux determination by closed-chamber methods can be seriously biased by inappropriate application of linear regression. *Biogeosciences*, *4*(6), 1005–1025. https://doi.org/10.5194/bg-4-1005-2007

Lohila, A., Aurela, M., Regina, K., & Laurila, T. (2003). Soil and total ecosystem respiration in agricultural fields: Effect of soil and crop type. *Plant and Soil*, *251*(2), 303–317. https://doi.org/10.1023/A:1023004205844

Martin, C. R., Zeng, N., Karion, A., Dickerson, R. R., Ren, X., Turpie, B. N., & Weber, K. J. (2017). Evaluation and environmental correction of ambient CO_2 measurements from a low-cost NDIR sensor. *Atmospheric Measurement Techniques*, *10*(7), 2383–2395. https://doi.org/10.5194/amt-10-2383-2017

Munson, S. M., Benton, T. J., Lauenroth, W. K., & Burke, I. C. (2010). Soil carbon flux following pulse precipitation events in the shortgrass steppe. *Ecological Research*, *25*(1), 205–211. https://doi.org/10.1007/s11284-009-0651-0

Ngatia, L. W., Reddy, K. R., Nair, P. K. R., Pringle, R. M., Palmer, T. M., & Turner, B. L. (2014). Seasonal patterns in decomposition and nutrient release from East African savanna grasses grown under contrasting nutrient conditions. *Agriculture, Ecosystems & Environment*, *188*, 12–19. https://doi.org/10.1016/j.agee.2014.02.004

Ohashi, M., Kume, T., Yamane, S., & Suzuki, M. (2007). Hot spots of soil respiration in an Asian tropical rainforest. *Geophysical Research Letters*, *34*(8), L08705. https://doi.org/10.1029/2007GL029587

Otieno, D. O., K'Otuto, G. O., Jákli, B., Schröttle, P., Maina, J. N., Jung, E., & Onyango, J. C. (2011). Spatial heterogeneity in ecosystem structure and productivity in a moist Kenyan savanna. *Plant Ecology*, *212*(5), 769–783. https://doi.org/10.1007/s11258-010-9863-1

Pantani, O.-L., Fioravanti, F., Stefanini, F. M., Berni, R., & Certini, G. (2020). Accounting for soil respiration variability – Case study in a Mediterranean pine-dominated forest. *Scientific Reports*, *10*(1), 1787. https://doi.org/10.1038/s41598-020-58664-6

Pellegrini, A. F. A., Hobbie, S. E., Reich, P. B., Jumpponen, A., Brookshire, E. N. J., Caprio, A. C., et al. (2020). Repeated fire shifts carbon and nitrogen cycling by changing plant inputs and soil decomposition across ecosystems. *Ecological Monographs*, *90*(4), e01409. https://doi.org/10.1002/ecm.1409

Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., et al. (2017). The value of soil respiration measurements for interpreting and modeling terrestrial carbon cycling. *Plant and Soil*, *413*(1), 1–25. https://doi.org/10.1007/s11104-016-3084-x

Poth, M., Anderson, I. C., Miranda, H. S., Miranda, A. C., & Riggan, P. J. (1995). The magnitude and persistence of soil NO, N_2O, CH_4, and CO_2 fluxes from burned tropical savanna in Brazil. *Global Biogeochemical Cycles*, *9*(4), 503–513. https://doi.org/10.1029/95GB02086

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Hari, P., Vesala, T., et al. (2004). Comparison of different chamber techniques for measuring soil CO_2 efflux. *Agricultural and Forest Meteorology*, *123*(3–4), 159–176. https://doi.org/10.1016/j.agrformet.2003.12.001

Rochette, P., Desjardins, R. L., & Pattey, E. (1991). Spatial and temporal variability of soil respiration in agricultural fields. *Canadian Journal of Soil Science*, *71*(2), 189–196. https://doi.org/10.4141/cjss91-018

Rodeghiero, M., & Cescatti, A. (2008). Spatial variability and optimal sampling strategy of soil respiration. *Forest Ecology and Management*, *255*(1), 106–112. https://doi.org/10.1016/j.foreco.2007.08.025

Savage, K. E., & Davidson, E. A. (2003). A comparison of manual and automated systems for soil CO_2 flux measurements: Trade-offs between spatial and temporal resolution. *Journal of Experimental Botany*, *54*(384), 891–899. https://doi.org/10.1093/jxb/erg121

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., et al. (2011). Persistence of soil organic matter as an ecosystem property. *Nature*, *478*(7367), 49–56. https://doi.org/10.1038/nature10386

Schmitz, O. J., Wilmers, C. C., Leroux, S. J., Doughty, C. E., Atwood, T. B., Galetti, M., et al. (2018). Animals and the zoogeochemistry of the carbon cycle. *Science*, *362*(6419), eaar3213. https://doi.org/10.1126/science.aar3213

Sitters, J., Kimuyu, D. M., Young, T. P., Claeys, P., & Olde Venterink, H. (2020). Negative effects of cattle on soil carbon and nutrient pools reversed by megaherbivores. *Nature Sustainability*, *3*(5), 360–366. https://doi.org/10.1038/s41893-020-0490-0

Somasundaram, J., Lal, R., Sinha, N. K., Dalal, R., Chitralekha, A., Chaudhary, R. S., & Patra, A. K. (2018). Cracks and potholes in vertisols: Characteristics, occurrence, and management. In *Advances in agronomy* (Vol. 149, pp. 93–159). Academic Press. https://doi.org/10.1016/bs.agron.2018.01.001

Stoyan, H., De-Polli, H., Böhm, S., Robertson, G. P., & Paul, E. A. (2000). Spatial heterogeneity of soil respiration and related properties at the plant scale. *Plant and Soil*, *222*(1), 203–214. https://doi.org/10.1023/A:1004757405147

Throop, H. L., Seely, M. K., & Marufu, V. J., & Summer Drylands Program Participants. (2020). Multiple scales of spatial heterogeneity control soil respiration responses to precipitation across a dryland rainfall gradient. *Plant and Soil*, *453*(1), 423–443. https://doi.org/10.1007/s11104-020-04614-0

Tuovinen, J.-P., Aurela, M., Hatakka, J., Räsänen, A., Virtanen, T., Mikola, J., et al. (2019). Interpreting eddy covariance data from heterogeneous Siberian tundra: Land-cover-specific methane fluxes and spatial representativeness. *Biogeosciences*, *16*(2), 255–274. https://doi.org/10.5194/bg-16-255-2019

Wachiye, S., Merbold, L., Vesala, T., Rinne, J., Räsänen, M., Leitner, S., & Pellikka, P. (2020). Soil greenhouse gas emissions under different land-use types in savanna ecosystems of Kenya. *Biogeosciences*, *17*(8), 2149–2167. https://doi.org/10.5194/bg-17-2149-2020

Wang, J., Li, Y., Bork, E. W., Richter, G. M., Chen, C., Shah, S. H. H., & Mezbahuddin, S. (2021). Effects of grazing management on spatio-temporal heterogeneity of soil carbon and greenhouse gas emissions of grasslands and rangelands: Monitoring, assessment and scaling-up. *Journal of Cleaner Production*, *288*, 125737. https://doi.org/10.1016/j.jclepro.2020.125737

Wen, Z., & Li, S. J. (2013). *Design and implementation of the greenhouse environmental monitoring system based on JN5148* (Vol. 333–335). Trans Tech Publications Ltd. https://doi.org/10.4028/www.scientific.net/AMM.333-335.2305

Xiao, J., Chen, J., Davis, K. J., & Reichstein, M. (2012). Advances in upscaling of eddy covariance measurements of carbon and water fluxes. *Journal of Geophysical Research*, *117*, G00J01. https://doi.org/10.1029/2011JG001889

Yasuda, T., Yonemura, S., & Tani, A. (2012). Comparison of the characteristics of small commercial NDIR CO_2 sensor models and development of a portable CO_2 measurement device. *Sensors*, *12*(3), 3641–3655. https://doi.org/10.3390/s120303641

Young, T. P., Porensky, L. M., Riginos, C., Veblen, K. E., Odadi, W. O., Kimuyu, D. M., et al. (2018). Relationships between cattle and biodiversity in multiuse landscape revealed by Kenya long-term exclosure experiment. *RAMA*, *71*(3), 281–291. https://doi.org/10.1016/j.rama.2018.01.005

Zhao, J.-F., Liao, Z.-Y., Yang, L.-Y., Shi, J.-K., & Tan, Z.-H. (2021). Characteristics of soil respiration and its components of a mixed dipterocarp forest in China. *Forests*, *12*(9), 1159. https://doi.org/10.3390/f12091159

References From the Supporting Information

Cozir-lp: Low power CO_2 sensor. (2022). Retrieved from https://cdn.shopify.com/s/files/1/0019/5952/files/CozIR-LP_Data_Sheet_Rev_4.7_0.pdf?v=1640012502

Pringle, R. M., & Tarnita, C. E. (2017). Spatial self-organization of ecosystems: Integrating multiple mechanisms of regular-pattern formation. *Annual Reviews*, *62*(1), 359–377. https://doi.org/10.1146/annurev-ento-031616-035413

Young, T. P., Okello, B. D., Kinyua, D., & Palmer, T. M. (1998). KLEE: A longterm multi-species herbivore exclusion experiment in Laikipia, Kenya. *African Journal of Range & Forage Science*, *14*(3), 94–102. https://doi.org/10.1080/10220119.1997.9647929