1 **The paradox of assessing greenhouse gases from soils for nature-**

2 **based solutions**

4 **Authors:**

5 Rodrigo Vargas[1*] and Van Huong Le[1]

7 **Affiliations**

8 [1]Department of Plant and Soil Science, University of Delaware, Newark, DE, USA

11 *Corresponding Author

12 Rodrigo Vargas (rvargas@udel.edu)

20   **Abstract**

21   Quantifying the role of soils in nature-based solutions requires accurate estimates of soil

22   greenhouse gas (GHG) fluxes. Technological advances allow us to measure multiple GHGs

23   simultaneously, and now it is possible to provide complete GHG budgets from soils (i.e.,

24   $CO_2$, $CH_4$, and $N_2O$ fluxes). We propose that there is a conflict between the convenience of

25   simultaneously measuring multiple soil GHG fluxes at fixed time intervals (e.g., once or

26   twice per month) and the intrinsic temporal variability and patterns of different GHG fluxes.

27   Information derived from fixed time intervals -commonly done during manual field

28   campaigns- had limitations to reproducing statistical properties, temporal dependence, annual

29   budgets, and associated uncertainty when compared with information derived from

30   continuous measurements (i.e., automated hourly measurements) for all soil GHG fluxes. We

31   present a novel approach (i.e., temporal univariate Latin Hypercube sampling) that can be

32   applied to provide insights and optimize monitoring efforts of GHG fluxes across time. We

33   suggest that multiple GHG fluxes should not be simultaneously measured at a few fixed time

34   intervals (mainly when measurements are limited to once per month), but an optimized

35   sampling approach can be used to reduce bias and uncertainty. These results have

36   implications for assessing GHG fluxes from soils and consequently reduce uncertainty on the

37   role of soils in nature-based solutions.

38

39   **Keywords:** Carbon dioxide, methane, nitrous oxide, representativeness, uncertainty

**1. Introduction**

Soils are essential for nature-based solutions for their role in climate mitigation potential through implementing different natural pathways (Griscom et al., 2017; Bossio et al., 2020). The climate mitigation potential of soils is dependent on multiple factors such as weather variability (Kim et al., 2012), ecosystem type (Oertel et al., 2016), soil structure (Ball, 2013), management practices (Shakoor et al., 2021), or disturbances (Vargas, 2012), where soils can ultimately act as net sources or sinks of greenhouse gases (GHGs). Therefore, accurate quantification of the magnitudes and patterns of soil GHGs fluxes is needed to understand the potential of soils to mitigate or contribute to global warming across ecosystems and different scenarios.

Most of our understanding of soil GHGs has come from manual measurements performed throughout labor-intensive field campaigns and experiments (Oertel et al., 2016). While most studies around the world have focused on soil $CO_2$ fluxes (Jian et al., 2020), early examples have reported coupled measurements of soil $CO_2$, $CH_4$, and $N_2O$ fluxes across tropical forests (Keller et al., 1986) and savannas (Hao et al., 1988), temperate forests (Bowden et al., 1993), and peatlands (Freeman et al., 1993). These pioneer studies provided an early view of the importance of integrated measurements of multiple soil GHG fluxes to understand the net global warming potential of soils but also demonstrated the technical limitations and challenges associated with these efforts. For example, it is known that manual measurements have the strength of providing good spatial coverage during field surveys but provide limited information about the temporal variability (Yao et al., 2009; Barba et al., 2021).

Technological advances have opened the opportunity to simultaneously measure multiple soil GHG fluxes (i.e., $CO_2$, $CH_4$, and $N_2O$) at unprecedented temporal resolution (e.g., hourly). These efforts have demonstrated differences in diel patterns and pulse events

65   (e.g., rewetting) due to wetting and drying cycles across tropical (Butterbach-Bahl et al.,

66   2004; Werner et al., 2007), subtropical (Rowlings et al., 2012), and temperate (Savage et al.,

67   2014; Petrakis et al., 2017) ecosystems. These approaches provide more accurate information

68   to calculate net GHG budgets and the global warming potential of soils (Capooci et al.,

69   2019). That said, performing automated measurements of multiple GHGs is expensive, and

70   this approach usually has a lower representation of the spatial heterogeneity within

71   ecosystems (Yao et al., 2009; Barba et al., 2021).

72        Ideally, we would like to measure everything, everywhere, and all the time, but this is

73   impossible due to logistical, technological, physical, and economic constraints. Lightweight

74   and low-powered laser-based spectrometers have reduced technical barriers to

75   simultaneously measuring multiple GHGs fluxes from soils. It is now easier and faster to

76   perform discrete manual surveys across time. This opportunity creates a paradox concerning

77   when to measure different GHG fluxes from soils when performing manual measurements.

78   Researchers generally tend to perform simultaneous measurements of multiple GHGs during

79   manual surveys, but this convenience could result in biased information. We propose that

80   there is a conflict between the convenience of measuring multiple GHGs at a few fixed time

81   intervals and the intrinsic temporal variability of magnitudes and patterns of different GHG

82   fluxes.

83        Here, we present a proof-of-concept and test how a subset of measurements derived

84   from a fixed temporal stratification (FTS) for simultaneous measurements (i.e., stratified

85   sampling schedule) or using an optimized sampling (i.e., temporal univariate Latin

86   Hypercube sampling ($tuLHs$)), compared with automated measurements of soil $CO_2$ ($F_A CO_2$),

87   $CH_4$ ($F_A CH_4$), and $N_2O$ ($F_A N_2O$) fluxes from a temperate forest (Petrakis et al., 2018; Barba

88   et al., 2021, 2019). The underlying assumption supporting any FTS approach is that a few

89   measurements in time can reproduce the statistical properties and temporal dependencies of

90  soil $CO_2$, $CH_4$, and $N_2O$ fluxes because these GHGs respond similarly to biological and

91  physical drivers. The *tuLHs* is a new statistical method for generating subsamples of

92  parameter values (i.e., soil GHG gas fluxes in this case study) to reproduce the probability

93  distribution and the temporal dependence of each original time series of GHG fluxes. We

94  reveal that reporting GHG fluxes using an FTS for simultaneous measurements may result in

95  biased information on temporal patterns and magnitudes. This study shows how a biased

96  sampling schedule could influence our understanding of GHG fluxes and, ultimately, the

97  climate mitigation potential of soils.

98

99  **2. Materials and Methods**

100  *2.1 Study site*

101  The experiment was performed in a temperate forest located at the St Jones Estuarine

102  Reserve (a component of the Delaware National Estuarine Research Reserve [DNERR] in

103  Delaware, USA. The site has a mean annual temperature of 13.3 °C and a mean annual

104  precipitation of 1119 mm. Soils are classified as Othello silt loam with a texture of 40% sand,

105  48% silt, and 12% clay within the first 10 cm (Petrakis et al., 2018). The dominant plant

106  species are bitternut hickory (*Carya cordiformis*), eastern red cedar (*Juniperus virginiana* L.),

107  American holly (*Ilex opaca*), sweet gum (*Liquidambar styraciflua* L.), and black gum (*Nyssa*

108  *sylvatica* (Marshall)). The site has a mean tree density of 678 stems ha$^{-1}$ and a diameter at

109  breast height (DBH) of 25.7±13.9 cm (mean±SD) (Barba et al., 2021).

110

111  *2.2 Automated measurements of soil GHG fluxes*

112  We analyzed data from automated measurements (1hr time intervals) of soil emissions of

113  three GHGs (i.e., $CO_2$, $CH_4$, and $N_2O$) between January and December 2015. This was a

114  typical year with a mean annual temperature of 13.4 °C and an annual precipitation of 1232

mm. Continuous measurements of soil GHGs were taken by coupling a closed-path infrared

gas analyzer (Li-COR LI-8100 A, Lincoln, Nebraska) and nine dynamic soil chambers (Li-

COR 8100–104) controlled by a multiplexer (Li-COR 8100-104) with a cavity ring-down

spectrometer (Picarro G2508, Santa Clara, California). A detailed description of the

experimental design and measurements protocol is described in previous studies (Petrakis et

al., 2018; Barba et al., 2021, 2019). Briefly, for each flux observation, we measured $CO_2$,

$CH_4$, and $N_2O$ concentrations every second with the Picarro G2508 for 300 seconds and

calculated fluxes (at 1 hr time intervals) from the mole dry fraction of each gas (i.e.,

corrected for water vapor dilution) using the SoilFluxPro software (v4.0; Li-COR, Lincoln,

Nebraska, USA). Fluxes were estimated using linear and exponential fits, and we kept the

flux calculation with the highest $R^2$. We applied quality assurance and quality control

protocols using information from all three GHGs as established in previous studies (Petrakis

et al., 2018; Barba et al., 2021, 2019; Capooci et al., 2019; Petrakis et al., 2017). Using these

time series, we extracted values to represent discrete temporal measurements based on FTS

and used the optimization approach described below.

*2.3 Temporal subsampling of time series*

Subsampling of time series was performed using FTS and a temporal optimization following

a univariate Latin Hypercube (*tuLHs*) approach. The difference between FTS and temporal

optimization is that the first approach is focused on a fixed schedule (e.g., sampling once per

month), and the second is focused on reproducing the statistical properties and temporal

dependence relationship of the original GHG time series with a subset of measurements. This

means optimized subsamples may not be spaced systematically (e.g., every 15 days), and

selected dates may vary for each GHG flux due to their specific statistical properties and

temporal dependence.

140    FTS represents a traditional schedule for performing manual measurements of GHG

141    fluxes from soils. The FTS is usually performed with manual measurements because they

142    require extensive logistical coordination due to travel time and costs, availability of

143    instrumentation (e.g., gas analyzers), personnel to perform the measurements, and weather

144    conditions. During these scheduled visits, researchers usually collect fluxes from all three

145    GHGs and analyze them systematically to calculate magnitudes and patterns throughout the

146    experiment. Usually, researchers perform manual samples during the early hours of the day

147    (between 9 am and 12 pm) to avoid confounding effects due to large changes in temperature

148    and moisture, as demonstrated by information summarized by the soil respiration global

149    database (Cueva et al., 2017; Jian et al., 2020). Consequently, we selected subsamples from

150    each original GHG time series (derived from automated measurements) using flux

151    measurements from 10 am at fixed intervals of once per month (n=12), twice per month

152    (n=24), or four times per month (n=48) starting on the first week of available data from

153    automated measurements.

154    We applied *tuLHs* as an alternative subsampling approach to obtain an optimized

155    subsample with the same univariate statistical properties and temporal dependence

156    relationship of the original GHG time series. Optimization was performed to select

157    subsamples for each GHG flux using the same number of samples as for FTS: twelve ($k$=12),

158    twenty-four ($k$=24), or forty-eight ($k$=48) measurements throughout the year of available data

159    from automated measurements.

160

161    *2.4 Temporal Univariate Latin Hypercube Sampling (tuLHs)*

162    Let $S = \{(x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_n, y_n, z_n)\}$ be observations of the variables *X, Y,* and *Z* in a

163    time series, where X, Y, and Z are soil GHGs (i.e., $CO_2$, $CH_4$, and $N_2O$). Each measured

164    variable is characterized by the univariate probability distribution function and the temporal

165    dependency function. Once these two functions are known, then the behaviors of the variable

166    can be reproduced (Le et al., 2020; Chilès and Delfiner, 2009; Trangmar et al., 1986; Pyrcz

167    and Deutsch, 2014). The *tuLHs* consists of three steps: (1) modeling the univariate behavior of

168    the variable using the empirical cumulative univariate probability distribution function; (2)

169    modeling the temporal dependence using the empirical variogram function; and (3) optimizing

170    a subsample applying a global optimization method, differential evolution, using the previously

171    obtained variogram function as an objective function.

172         First, to model the univariate behavior of the variables from the observations of S, the

173    empirical univariate cumulative distribution function $F^*_n$ (x) of X is estimated by:

174
$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I\{x_i \leq x\} \quad (1)$$

175    where *I* represents an indicator function equal to 1 when its argument is true, and 0 otherwise.

176    Similarly, the empirical univariate distribution function of the variables *Y* and *Z* can be derived.

177    Second, to model the temporal dependence of the variables from the observations of *S*, the

178    empirical temporal correlation function (i.e., temporal variogram function) $\gamma^*(t)$ of X is

179    estimated by:

180
$$\gamma^*(t) = \frac{1}{2N(t)}\sum_{i=1}^{N(t)} [X(t_i + t) - X(t_i)]^2 \quad (2)$$

181    where *N(t)* is the number of pairs $X(t_i + t)$ $and$ $X(t_i)$ are separated by a time *t*. The variogram

182    functions of the variables Y and Z are analogous. Third, to optimize the subsample, it is

183    required to choose the "optimal" data points with the selected sample size (i.e., *k*=12, 24, or

184    48; where $k \ll n)$ that will have the same behavior as the original observations of *S* (i.e., GHG

185    fluxes derived from automated measurements). To achieve this objective, we use differential

186    evolution, a global optimization method (Storn and Price, 1997), using the variogram function

187    as an objective function. The procedure consists of dividing the univariate empirical probability

188    distribution in Eq. (1) into *k* equiprobable strata, which is equivalent to k ordered data subsets.

189    From each subset, only one value must be chosen to satisfy the condition of a univariate Latin

190    hypercube. The differential evolution method is applied to find the optimal points that

191    minimize the difference between the subsample variogram γ (t) and the data variogram γ * (t)

192    in Eq. (3).

193
$$OF_1 = \sum_{i=1}^{N(t)} [\gamma(t) - \gamma^*(t)]^2 \qquad (3)$$

194    where *OF* is the objective function, and the variograms γ (t) and γ * (t) are calculated using Eq.

195    (2).

196

197    *2.5 Statistical analyses*

198    The t-test was used to compare the means, and the Kolmogorov-Smirnov test was used to

199    compare the probability distribution of measurements derived from each sampling protocol.

200    All tests were done with a 95% confidence level. In addition, their statistical properties, such

201    as mean, median, standard deviation, and first and third quartile, were compared. The

202    differences in the experimental semivariograms were calculated as a comparison measure for

203    the temporal dependence of the samples and the original time series of GHG fluxes. For

204    cumulative sums of GHG flux, their mean is calculated as the most likely value, and their

205    quantile difference between 97.5 and 2.5 is used to quantify the range of uncertainty. All

206    analyzes were performed using the R program (Team and Others, 2013).

207

208    **3. Results**

209    *3.1 Relationships among GHG fluxes from soils*

210    Justification in support of FTS for simultaneous measurements of GHG fluxes would require

211    evidence of strong linear correlations between magnitudes and temporal dependence among

soil GHG fluxes. First, we did not find strong linear relationships between any combination of GHG fluxes from soils derived from automated measurements (Fig. S1). Therefore, our data did not support the assumption that the magnitude of one GHG flux was associated with a linear increase or decrease of another GHG flux. Second, semivariogram models demonstrated differences in the temporal dependence for each GHG flux. Automated measurements of soil $CO_2$ fluxes ($F_ACO_2$) showed a temporal dependence following a Gaussian variogram model, with a nugget of 4, a sill plus nugget of 28, and a correlation range of 80 days (Fig. S2a). Automated measurements of soil $CH_4$ fluxes ($F_ACH_4$) also showed a temporal dependence but followed a spherical variogram model, with a nugget of $7 \times 10^{-8}$, a sill plus nugget of $1.5 \times 10^{-7}$, and a correlation range of 110 days (Fig. S2b). In contrast, automated measurements of soil $N_2O$ fluxes ($F_AN_2O$) did not show a temporal dependence, where a pure nugget effect was present, and with a correlation range of 0 days (Fig. S2c). Consequently, these GHG fluxes' magnitudes and temporal patterns were different and did not support FTS for simultaneous measurements of GHG from soils.

*3.2 Optimization of GHG sampling protocols*

We applied a *tuLHs* approach to identify subsamples with the same statistical properties and temporal dependence for each of the original GHG time series from automated measurements. Subsamples were identified for twelve ($k=12$), twenty-four ($k=24$), or forty-eight ($k=48$) measurements throughout the year for each GHG time series. Our results show that the optimized measurement dates were different for each GHG flux (Fig. 1), and we provide explicit examples for $k=24$ (Fig. 1) and $k=12, 48$ (Fig. S3, S4).

The optimized $CO_2$ subsamples were well distributed throughout the year for all sampling scenarios (i.e., $k$ from 12 to 48) because $F_ACO_2$ had a strong temporal dependence and a small nugget effect with respect to the sill (Fig. S2a). The optimized $CH_4$ subsamples

237    were also relatively well distributed throughout the year, especially for scenarios of $k$=24 and

238    $k$=48, as $F_A CH_4$ also had a temporal dependence but with a higher nugget effect with respect

239    to the sill (Fig. S2b). Finally, the optimized $N_2O$ subsamples were more challenging to

240    define, especially with a small sample size (i.e., $k$=12; Fig. S3c) because $F_A N_2O$ did not have

241    a temporal dependence (Fig. S2c).

242

243    *3.3 Differences in statistical properties and temporal dependency of subsamples*

244    Overall, there were no statistically significant differences between the mean values derived

245    from automated measurements and those from FTS or the *tuLHs* approach (Fig. 2 for $k$=24;

246    Fig. S5 for $k$=12; Fig. S6 for $k$=48; Tables S1 and S2). Although this appears promising, more

247    than a simple comparison of the means is needed to evaluate the information derived from

248    different sampling approaches. In other words, it is possible to have a similar mean value

249    without reproducing the probability distribution nor the temporal dependence of the original

250    time series (i.e., correct answer but for the wrong reasons). Here, we present results based on

251    comparing the means, standard deviation, probability distributions, and semivariograms

252    derived from automated measurements and the different sampling scenarios for all GHG

253    fluxes.

254         The mean of $F_A CO_2$ was 5.9 µmol $CO_2$ m$^{-2}$ s$^{-1}$, while the mean for FTS 5.5 µmol $CO_2$

255    m$^{-2}$ s$^{-1}$ and 5.9 µmol $CO_2$ m$^{-2}$ s$^{-1}$ for the *tuLHs* approach with $k$=24 (Fig. 3a-c). These results

256    were comparable with the means derived from FTS (5.4 and 5.4 µmol $CO_2$ m$^{-2}$ s$^{-1}$) and the

257    *tuLHs* approach (6.2 and 5.9 µmol $CO_2$ m$^{-2}$ s$^{-1}$) using $k$=12 and $k$=48, respectively (Figs. S5,

258    S6; Table S1). The standard deviation of $F_A CO_2$ was 3.9 and 3.2 µmol $CO_2$ m$^{-2}$ s$^{-1}$ for FTS,

259    and 3.9 µmol $CO_2$ m$^{-2}$ s$^{-1}$ for the *tuLHs* approach with $k$=24 (Figs. 3a-c). These results were

260    comparable with the standard deviations derived from FTS (3.1 and 3.3 µmol $CO_2$ m$^{-2}$ s$^{-1}$) and

261    the *tuLHs* approach (4.1 and 3.9 µmol $CO_2$ m$^{-2}$ s$^{-1}$) using $k$=12 and $k$=48, respectively (Fig. S5,

262   S6; Table S1). Our results show that the semivariograms of optimized samples using the *tuLHs*

263   approach closely approximate the semivariograms of automated measurements for $k=24$ (Fig.

264   4a) and $k=12$ and 48 (Figs. S7a, S8a). These results are consistent with the sums of absolute

265   differences between the semivariograms of the samples and the semivariogram of $F_ACO_2$ with

266   differences of 69.31, 54.39, 49.42 for FTS, and 5.69, 1.99, 1.39 for the *tuLHs* approach for

267   $k=12$, 24, 48, respectively (Table S2).

268          The mean of $F_ACH_4$ was -0.93 nmol $CH_4$ $m^{-2}$ $s^{-1}$, while -0.86 nmol $CH_4$ $m^{-2}$ $s^{-1}$ for FTS

269   and -0.94 nmol $CH_4$ $m^{-2}$ $s^{-1}$ for the *tuLHs* approach with $k=24$ (Fig. 3d-f). These results were

270   also comparable with the means derived from FTS (-0.83 and -0.88 nmol $CH_4$ $m^{-2}$ $s^{-1}$) and the

271   *tuLHs* approach (-0.87 and -0.92 nmol $CH_4$ $m^{-2}$ $s^{-1}$) using $k=12$ and 48, respectively (Figs. S5,

272   S6; Table S1). The standard deviation of $F_ACH_4$ was 0.36 and 0.26 nmol $CH_4$ $m^{-2}$ $s^{-1}$ for FTS

273   and 0.34 nmol $CH_4$ $m^{-2}$ $s^{-1}$ for the *tuLHs* approach with $k=24$. These results were comparable

274   with the standard deviations derived from FTS (0.27 and 0.29 nmol $CH_4$ $m^{-2}$ $s^{-1}$) and the *tuLHs*

275   approach (0.33 and 0.35 nmol $CH_4$ $m^{-2}$ $s^{-1}$) using $k=12$ and $k=48$, respectively (Figs. S5, S6;

276   Table S1). The semivariograms of optimized samples using the *tuLHs* approach closely

277   approximate the semivariogram of automated measurements for $k=24$ (Fig. 4b) and $k=12$ and

278   48 (Figs. S7b, S8b). Consequently, the sums of absolute differences between the

279   semivariograms of the samples and the semivariogram of $F_ACH_4$ were 0.63, 0.48, 0.49 for FTS,

280   and 0.06, 0.04, 0.02 for the *tuLHs* approach with $k=12$, 24, 48, respectively (Table S2).

281          Finally, the mean of $F_AN_2O$ was 0.45 and 0.61 nmol $N_2O$ $m^{-2}$ $s^{-1}$ for FTS, and 0.51 nmol

282   $N_2O$ $m^{-2}$ $s^{-1}$ for the *tuLHs* approach with $k=24$ (Fig. 3g-i). These results were also comparable

283   with the means derived from FTS (0.59 and 0.25 nmol $N_2O$ $m^{-2}$ $s^{-1}$) and the *tuLHs* approach

284   (0.58 and 0.49 nmol $N_2O$ $m^{-2}$ $s^{-1}$) using $k=12$ and 48, respectively (Figs. S5, S6; Table S1). The

285   standard deviation of $F_AN_2O$ was 1.62 and 1.97 nmol $N_2O$ $m^{-2}$ $s^{-1}$ for FTS, and 1.54 nmol $N_2O$

286   $m^{-2}$ $s^{-1}$ for the *tuLHs* approach with $k=24$. These results were comparable with the standard

287     deviations derived from FTS (1.38 and 0.91 nmol $N_2O$ $m^{-2}$ $s^{-1}$) and the *tuLHs* approach (1.58

288     and 1.54 nmol $N_2O$ $m^{-2}$ $s^{-1}$) using $k=12$ and $k=48$, respectively (Figs. S5, S6; Table S1). Our

289     results show no temporal dependence for $N_2O$ fluxes, but the semivariograms of optimized

290     samples using the *tuLHs* approach closely approximate the semivariogram of automated

291     measurements for $k=24$ (Fig. 4c) and $k=12$ and 48 (Figs. S7c, S8c). Consistently, the sum of

292     absolute differences between the semivariograms of the samples and the semivariogram of

293     $F_A N_2O$ were 10.01, 12.25, 16.75 for FTS, and 0.82, 1.13, 3.57 for the *tuLHs* approach with

294     $k=12$, 24, 48, respectively (Table S2).

295        These results show that the *tuLHs* approach reproduced the probability distribution and

296     the temporal dependence of the time series derived from automated measurements with more

297     precision than FTS for all GHGs. In the next section, we explore the implications of these

298     differences for calculating cumulative GHG fluxes.

299

300     *3.4 Calculation of cumulative GHG fluxes*

301     We calculated the cumulative flux for all GHGs using available information from automated

302     measurements (Fig. 2; Table S3). The cumulative sum for available measurements of $F_A CO_2$

303     was 5758.5 g $CO_2$ $m^{-2}$ [893.9, 13860.8; 95% CI]; for $F_A CH_4$ was -0.47 g $CH_4$ $m^{-2}$ [-0.81, -

304     0.19; 95% CI]; and 0.63 g $N_2O$ $m^{-2}$ [-0.75, 5.19; 95% CI] for $F_A N_2O$.

305        We used the mean for each GHG flux derived from the *tuLHs* approach or the FTS to

306     calculate the cumulative sum (Table S3). We found that the FTS underestimated the

307     cumulative flux (-8.4, -6.2, -7.1%) and the uncertainty (-32.6, -21.6, -19.3%) of $F_A CO_2$ for

308     $k=12$, 24, 48, respectively (Fig. 5a). In contrast, the *tuLHs* approach slightly overestimated

309     the cumulative flux (6.5, 1.1, 0.1%) and slightly underestimated the uncertainty (-9.1, -4.4, -

310     3.7%) for $k=12$, 24, 48, respectively (Fig. 5a).

311       The FTS underestimated the cumulative flux (-9.1, -6.1, -3.1%) and the uncertainty (-

312    31.8, -27.3, -15.9%) of $F_A CH_4$ for $k$=12, 24, 48, respectively (Fig. 5b). In contrast, the *tuLHs*

313    approach underestimated the cumulative flux (-6.1%) only for $k$=12, but slightly

314    underestimated the uncertainty (-15.9, -6.8, -4.5%) for $k$=12, 24, 48, respectively (Fig. 5b).

315       The FTS substantially underestimated the cumulative flux (-168, -170, -173%) of

316    $F_A N_2O$ for $k$=12, 24, 48, respectively. Uncertainty was overestimated for $k$=12 and 24 (3.6

317    and 26%) and underestimated for $k$=48 (-31%; Fig. 5c). In contrast, the *tuLHs* approach

318    overestimated less the cumulative flux (29.5, 13.4, 9.1%) for $k$=12, 24, 48, respectively (Fig.

319    5c). This approach underestimated the uncertainty for $k$=12 (-11.2%) and $k$=24 (-13.8%) but

320    overestimated the uncertainty by 2.9% for $k$=48 (Fig. 5c). These results show that the *tuLHs*

321    approach consistently provided closer estimates for cumulative sums and uncertainty ranges

322    than an FTS for all GHG fluxes.

323

324    **4. Discussion**

325    Applied challenges, such as quantifying the role of soils in nature-based solutions, require

326    accurate estimates of GHG fluxes. To do this, two fundamental problems exist for designing

327    environmental monitoring protocols: where and when to measure? Ultimately a monitoring

328    protocol aims to quantify the attributes of an ecosystem so that it can be compared in time

329    within that ecosystem or with other ecosystems. Because we cannot measure everything,

330    everywhere, and all the time, we can argue that any monitoring protocol has assumptions

331    based on physical, economic, social, and practical reasons to address a specific scientific

332    question. These assumptions for designing monitoring protocols could result in misleading,

333    biased, or wrong conclusions, and therefore is critical to assess the consequences of different

334    monitoring efforts. As Hutchinson described in "The Concept of Pattern in Ecology," we do

335     not always know if a given pattern is extraordinary or a simple expression of something

336     which we may learn to expect all the time (Hutchinson, 1953).

337         Automated measurements have revolutionized our understanding of the temporal

338     patterns and magnitudes of soil GHG fluxes (Savage et al., 2014; Bond-Lamberty et al.,

339     2020; Tang et al., 2006; Capooci and Vargas, 2022b). These measurements have limitations

340     in representing spatial variability and have higher equipment costs that limit their broad

341     applicability across study sites (Vargas et al., 2011). Consequently, discrete manual

342     measurements are a common approach to simultaneously measure multiple GHG fluxes and

343     report patterns, budgets, and information to parameterize empirical and process-based models

344     (Phillips et al., 2017; Wang and Chen, 2012). In this study, we argue that the convenience of

345     simultaneously measuring multiple GHGs using FTS may result in biased estimates.

346     Therefore, optimization of sampling protocols is needed to provide insights to improve

347     measurement protocols when there is a limited number of measurements in time (i.e., $k$=12,

348     24, 48).

349         We show that the magnitude of one GHG flux is not associated with a linear increase

350     or decrease of another GHG flux, and the temporal dependencies of each GHG flux are

351     different (Fig. S1). Therefore, it is not possible to infer the dynamics of one GHG flux based

352     solely on information from another under the assumption that they share similar (or

353     autocorrelated) biophysical drivers. These results imply that the magnitudes and temporal

354     patterns of GHGs are different and therefore do not support an FTS approach for

355     simultaneous measurements of GHG fluxes from soils.

356         Multiple studies have shown that the relevance of different biophysical drivers (e.g.,

357     temperature, moisture, light) is different for soil $CO_2$, $CH_4$, or $N_2O$ fluxes (Luo et al., 2013;

358     Tang et al., 2006; Ojanen et al., 2010). Our results show that soil $CO_2$ fluxes have a strong

359     temporal dependence (Fig. S2a), likely due to the strong relationship between these fluxes

360    and soil temperature in this and other temperate mesic ecosystems (Hill et al., 2021; Bahn et

361    al., 2010; Barba et al., 2019). The temporal dependence decreased for soil $CH_4$ fluxes (Fig.

362    S2b), where there is less evidence for such a strong correlation with soil temperature in this

363    and other temperate mesic ecosystems (Bowden et al., 1998; Castro et al., 1995; Warner et

364    al., 2019; Barba et al., 2019). It has been reported that multiple variables and complex

365    relationships are usually needed to explain the variability of soil $CH_4$ fluxes in forest soils, as

366    there is a delicate balance between methanogenesis and methanotrophy (Luo et al., 2013;

367    Castro et al., 1994; Murguia-Flores et al., 2018). In contrast, soil $N_2O$ fluxes had no temporal

368    dependence (Fig. S2c), showing decoupling from the observed patterns of soil $CO_2$ and $CH_4$

369    fluxes (Wu et al., 2010), likely as a result of independent biophysical drivers regulating soil

370    $N_2O$ fluxes  (Luo et al., 2013; Bowden et al., 1993; Ullah and Moore, 2011).

371        To address the limitations of an FTS protocol, we propose a novel optimization

372    approach (i.e., *tuLHs*) to reproduce the probability distribution and the temporal dependence

373    of each original time series of GHG fluxes. Traditional methods usually optimize subsamples

374    by either individually focusing on reproducing the probability distribution of the original

375    information (Huntington and Lyrintzis, 1998) or reproducing the temporal dependence of the

376    original information (Gunawardana et al., 2011). The *tuLHs* is a simple approach that uses

377    the univariate probability distribution function and the temporal correlation function (i.e.,

378    variogram) as objective functions for each GHG flux. Our results show that optimized

379    subsamples do not coincide in time for the three GHGs, suggesting that information should

380    be collected based on each GHG flux's specific statistical and temporal characteristics (Fig.

381    1). This study provides proof-of-concept for the application of the *tuLHs*. It demonstrates

382    how an optimization approach provides insights to design monitoring protocols and improve

383    soil GHG flux estimates.

384       The more temporal data we can collect, the better, but in many cases, measurement

385    protocols are limited to a few measurements per year (i.e., $k$=12 to 48). Our results

386    demonstrate that for a small sample size (i.e., $k$=12), the optimized measurements for soil

387    $CO_2$ fluxes are consistently spread across the year, and for soil $CH_4$ fluxes are centered

388    within the growing season because of their strong temporal dependence. For the case of soil

389    $N_2O$ fluxes, the variogram shows a constant temporal variability, meaning there is no

390    temporal dependence. Therefore, the optimized measurements are concentrated within the

391    fall season due to their distribution probability (Fig. 1a). Our optimization approach shows

392    how measurements can be distributed across time as more samples are available (i.e., $k$=24 to

393    48; Fig. 1b-c) and demonstrates that optimization is critical when a limited number of

394    measurements are available. In other words, a few measurements properly distributed across

395    time provide better agreement with information derived from automated measurements. A

396    similar conclusion was proposed for the spatial distribution of environmental observatory

397    networks, where a network of few sites properly distributed (e.g., across a country) improves

398    our understanding of the target variable more than a spatially biased network (Villarreal et

399    al., 2019). Thus, the representativeness assessment of information collected across time and

400    space is needed to evaluate environmental measurements and quantify nature-based solutions

401    accurately.

402       We highlight that this optimization approach should be implemented across different

403    ecosystems as it will result in site-specific recommendations. The *tuLHs* can be applied to

404    any time series length and with any time step (e.g., hours, days), but specific results will be

405    representative of the probability distribution and the temporal dependence of the selected

406    time series. What is essential is to question if a few measurements from an experiment

407    represent the *reality* of the physical world because if limited information is available, then the

408    actual probability distribution and temporal dependence of the phenomena could be an

409    unknown-unknown. In other words, with few measurements, we may not be aware, and we

410    will not be able to know which is the actual probability distribution and temporal dependence

411    of the studied phenomena. To address this challenge, we tested the tuLHs approach with

412    high-temporal frequency information representing the probability distribution of multiple soil

413    GHG fluxes at the daily time-step across a calendar year.

414        In this case study, the year chosen had typical climatological conditions and

415    demonstrates that the statistical properties of the GHG fluxes are different and do not support

416    an FTS approach. Therefore, longer time series (e.g., multi-year) may provide more robust

417    optimizations that can be applied to monitoring efforts in future years. Alternatively, forecast

418    scenarios can be predicted, or in some cases (e.g., soil $CO_2$ efflux in temperate mesic

419    ecosystems), a proxy variable could be used (e.g., soil temperature) to inform the *tuLHs* and

420    provide insights for a sampling design. In those cases, the *tuLHs* can be used to suggest an

421    optimized sampling design under those specific assumptions. Ideally, automated

422    measurements should be co-located with manual efforts to adequately capture the temporal

423    and spatial variability of soil GHG fluxes at a specific site.

424        There are several implications of biased monitoring protocols for understanding soil

425    GHG fluxes and nature-based solutions. First, temporal patterns and temporal dependency

426    may need to be revisited for studies using an FTS approach. Soil GHG fluxes have complex

427    temporal dynamics that vary from diurnal to seasonal and annual scales (Vargas et al., 2010)

428    that a few measurements following an FTS approach cannot reproduce (Barba et al., 2019;

429    Bréchet et al., 2021). Second, soil GHG fluxes could present hot-moments, which are

430    transient events with disproportionately high values that are often missed with an FTS

431    approach (Vargas et al., 2018; Butterbach-Bahl et al., 2004). Third, cumulative sums and

432    uncertainty ranges are biased or misleading when derived using an FTS approach (Tallec et

433    al., 2019; Lucas-Moffat et al., 2018; Capooci and Vargas, 2022b). Our study demonstrates

434     that an optimized approach consistently provided closer estimates for cumulative sums and

435     uncertainty ranges when compared with automated measurements (Fig. 5). We postulate that

436     representing the variability of soil $N_2O$ fluxes is more sensitive to the FTS approach (>170%

437     and >30% for cumulative sums and uncertainty ranges, respectively) than for soil $CH_4$ and

438     $CO_2$ fluxes. Fourth, it is possible that if the information derived from an FTS approach is

439     biased, then functional relationships could also be different from those derived from

440     automated measurements (Capooci and Vargas, 2022a). It has been argued that hypothesis

441     testing and our capability of forecasting responses of soil GHG fluxes to changing climate

442     conditions is also biased with information from the FTS approach (Vicca et al., 2014).

443     Finally, because soils have a central role in nature-based solutions within countries and

444     across the world (Griscom et al., 2017; Bossio et al., 2020), accurate measurements are

445     required to assess management practices, environmental variability, and the contribution of

446     GHGs from soils.

447

448     **Conclusion**

449     We highlight that we only sometimes know if a given pattern is extraordinary or a simple

450     expression of something which we may learn to expect all the time (Hutchinson, 1953).

451     Arguably, there is bias in our understanding of the probability distribution and temporal

452     dependency of soil GHG fluxes across the world because most results are based on a few

453     manual measurements (e.g., once a month) following an FTS approach. Currently, it is

454     unknown how large such bias could be across studies and ecosystems, but because most

455     studies lack high-temporal frequency information, the real probability distribution and

456     temporal dependency of soil GHG fluxes may remain unknown in most study sites. What is

457     essential is to question if the observed patterns, derived from an FTS approach, are enough

458    for improving our understanding of soil processes or are results that we have learned to

459    expect.

460          We postulate that with emergent technologies, there is a convenience of measuring

461    multiple GHGs from soils; however, few measurements collected at fixed time intervals

462    result in biased estimates. We recognize that potential measurement bias depends on each

463    GHG flux's magnitudes and temporal patterns and could be site-specific. Nevertheless,

464    evaluations are needed to quantify potential bias in estimates of GHG budgets and

465    information used for model parameterization and environmental assessments. Furthermore,

466    the underlying assumption that each GHG flux responds similarly to biophysical drivers may

467    need to be tested across multiple ecosystems to quantify how few measurements influence

468    our understanding of magnitudes and temporal patterns of soil GHG fluxes.

469          In this study, we present a proof-of-concept and propose a novel approach (i.e.,

470    temporal univariate Latin Hypercube sampling) that can be applied with site-specific

471    information of different ecosystems to improve monitoring efforts and reduce the bias of

472    GHG flux measurements across time. We highlight that constant biased environmental

473    monitoring may provide confirmatory information, which we have learned to expect, but

474    modifications of monitoring protocols could shed light on new or unexpected patterns. These

475    new patterns are the ones that will test paradigms and push science frontiers.

476

477

478    *Data Availability.* All data used for this analysis is available at:

479    https://doi.org/10.6084/m9.figshare.19556004.v1. The R code used in this study is available

480    at: https://github.com/vargaslab/temporal-univariate-Latin-Hypercube.git

481

482

483 *Author Contributions.* R.V. conceived this study, and V.H.L. designed and performed the

484 primary analysis with input from R.V in all phases. R.V. wrote the manuscript with

485 contributions from V.H.L.

486

487 *Competing Interest Statement.* None

488

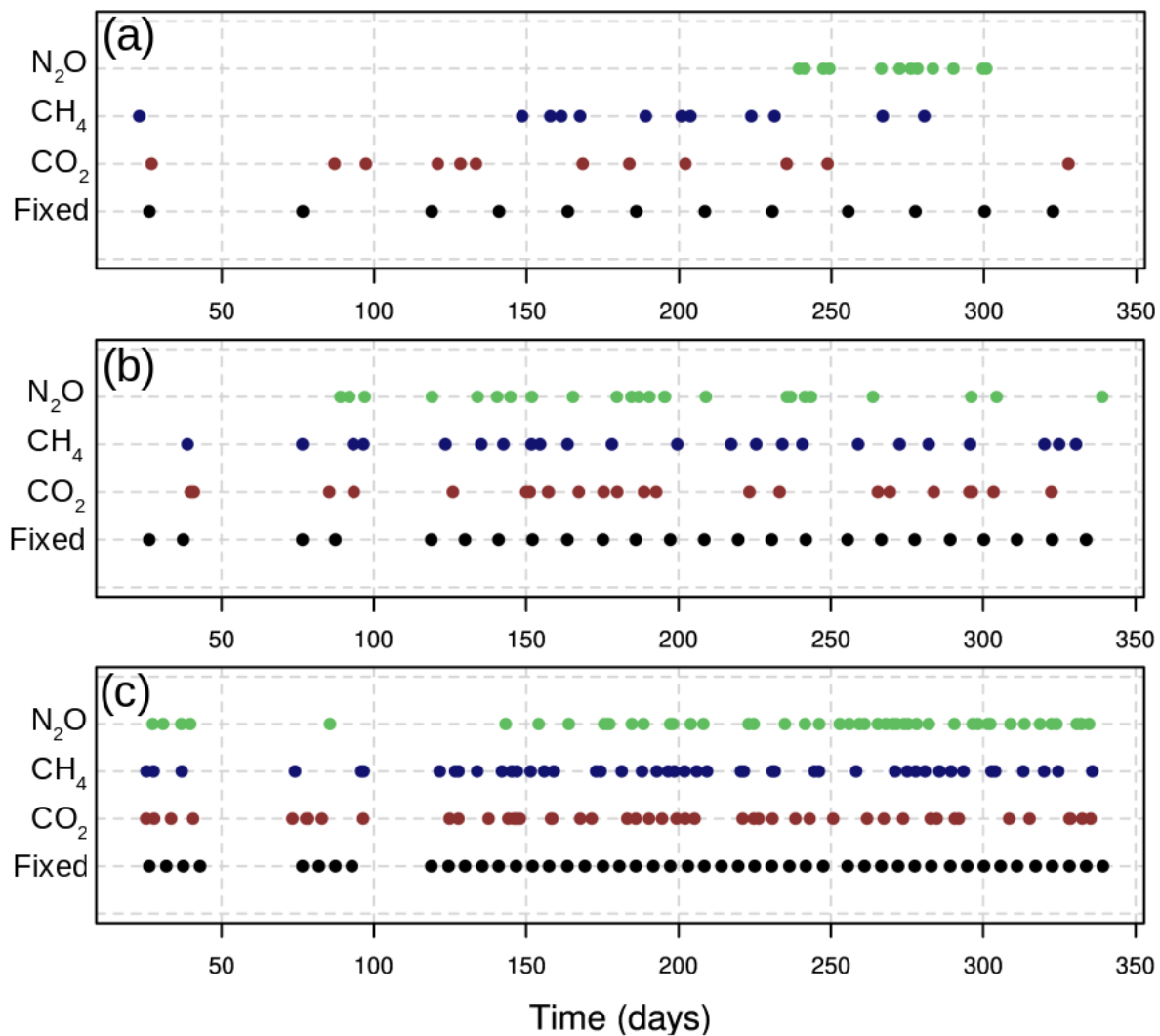
489 *Acknowledgments.* The authors thank the Delaware National Estuarine Research Reserve

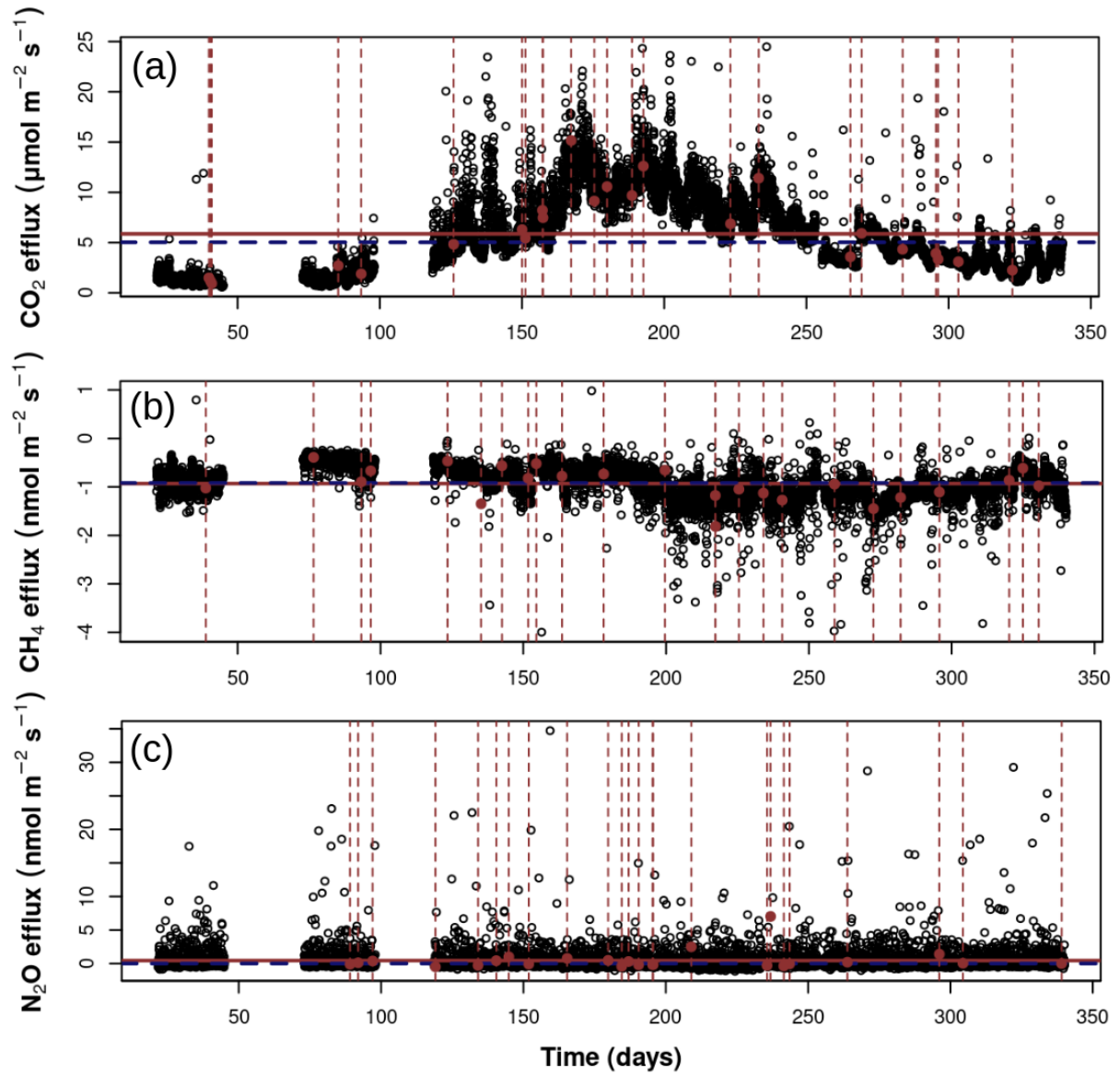
490 (DNERR) and the St Jones Reserve personnel for their support throughout this study. The

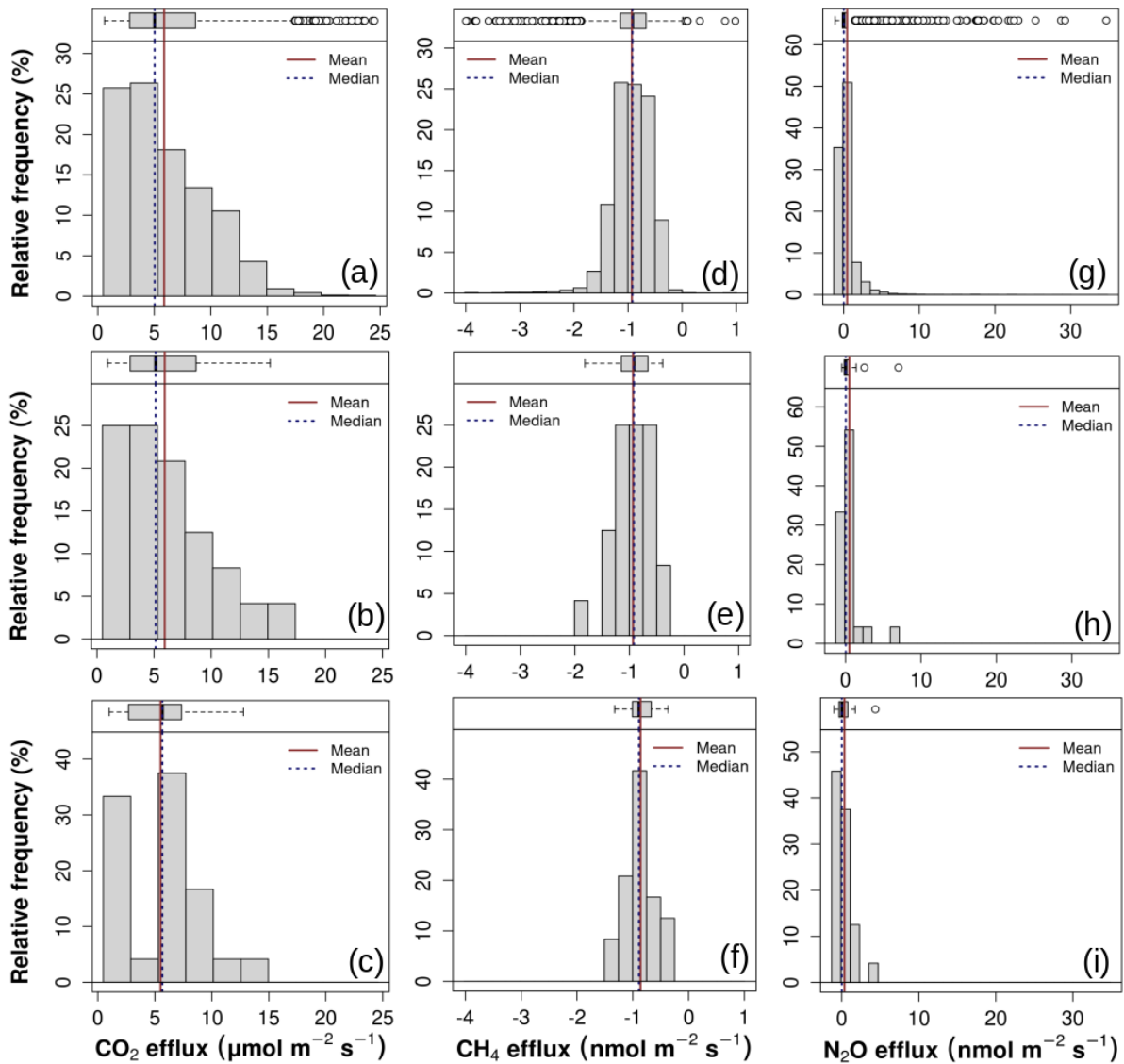
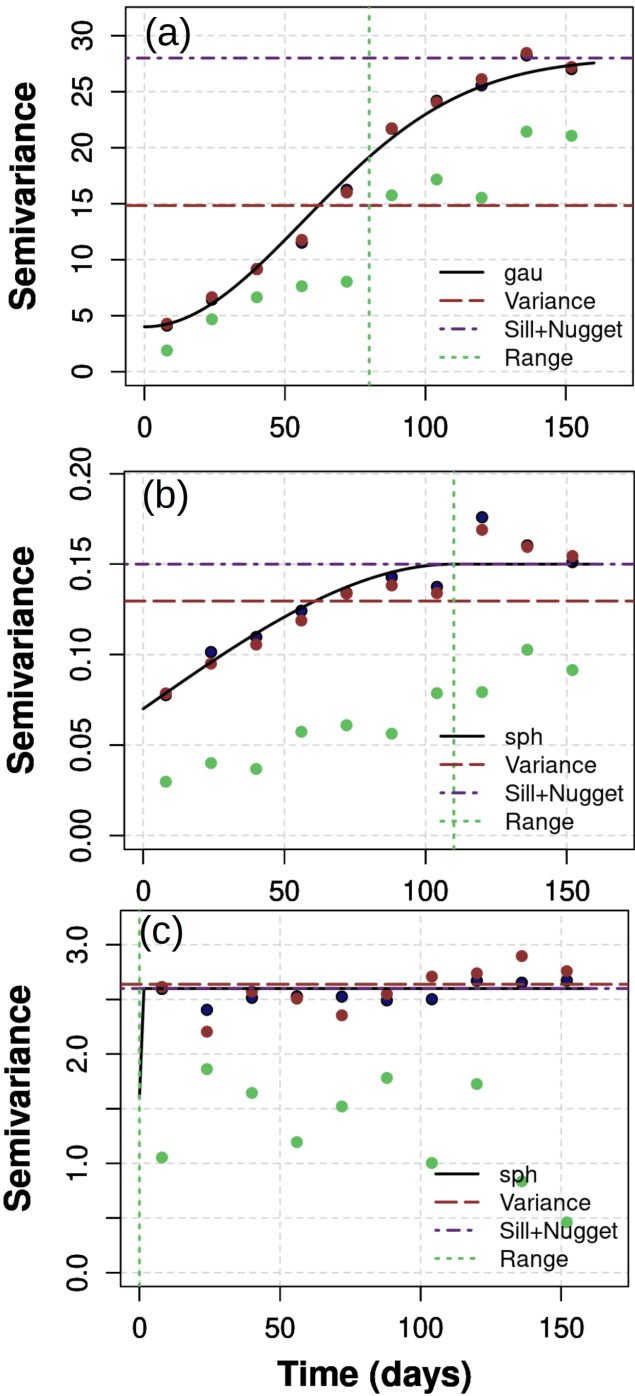
491 authors acknowledge the land on which they realized this study as the traditional home of the

492 Lenni-Lenape tribal nation (Delaware nation). This study was funded by a grant from the

493 National Science Foundation (#1652594) and NASA Carbon Monitoring System

494 (80NSSC21K0964).

498
499
500
501
502 **Figure 1.** Temporal distribution of fixed temporal stratification (i.e., stratified manual
503 sampling approach) and optimized sampling using a temporal univariate Latin Hypercube
504 (*tuLHs*) approach for: $k$=12 (a), $k$=24 (b), and $k$=48 (c). Fixed temporal stratification is in
505 black, soil $CO_2$ fluxes in red, soil $CH_4$ fluxes in blue, and soil $N_2O$ fluxes in green. Time (x-
506 axis) represents days from January 1 to December 31 of, 2015.

507

508
509
510 **Figure 2.** Time series of automated measurements ($F_A$) of soil greenhouse gas fluxes (black
511 circles) and optimized samples ($k$=24) using a temporal univariate Latin Hypercube sampling
512 (*tuLHs*) approach for soil $CO_2$ (a), soil $CH_4$ (b) and soil $N_2O$ (c) fluxes. The horizontal red
513 line represents the mean, and the horizontal blue line is the median of each greenhouse gas
514 flux derived from automated measurements. The selection of data points for $k$=12 and 48 are
515 presented for each soil greenhouse gas time series in Figs. S3 and S4, respectively. Time (x-
516 axis) represents days from January 1 to December 31 of, 2015.
517
518

519

520
521
522
523 **Figure 3**. Histograms for automated measurements of soil $CO_2$ ($F_A$ $CO_2$; a), soil $CH_4$ ($F_A$
524 $CH_4$; d), and soil $N_2O$ ($F_A$ $N_2O$; g). Histograms for optimized samples ($k$=24) using a
525 temporal univariate Latin Hypercube sampling (*tuLHs*) approach for soil $CO_2$ (b), soil $CH_4$
526 (e), and soil $N_2O$ (h) fluxes. Histograms for fixed temporal stratification (i.e., stratified
527 manual sampling schedule) ($k$=24) for soil $CO_2$ (c), soil $CH_4$ (f), and soil $N_2O$ (i) fluxes.
528 Supplementary material includes results for measurements with $k$=12 (Fig. S5) and $k$=48
529 (Fig. S6).
530
531

533
534
535
536 **Figure 4**. Comparison of semivariograms between automated measurements ($F_A$) of soil
537 greenhouse gas fluxes (solid black line) and for optimized samples using a temporal
538 univariate Latin Hypercube sampling (*tuLHs*) approach (red circles) or fixed temporal
539 stratification (green circles) with $k$=24. Semivarograms are presented for soil $CO_2$ (a), $CH_4$
540 (d), and $N_2O$ (c) fluxes. Semivariograms for measurements with $k$=12 and $k$=48 are shown in
541 Supplementary Figs. S7 and S8, respectively. Semivariogram fits were gaussian (Gau) or
542 spherical (sph).

543
544

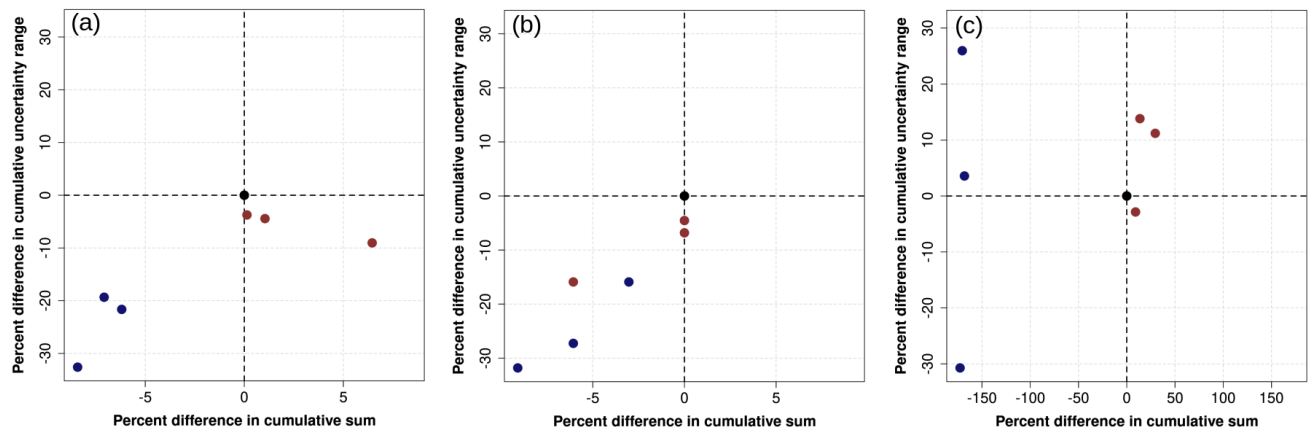



**Figure 5.** Comparison of percent differences from cumulative sums and associated uncertainty (95% CI) between greenhouse gas fluxes derived from automated measurements ($F_A$) and using an optimized sampling approach (*tuLHs*) or a fixed temporal stratification. Differences are represented for soil $CO_2$ (a), soil $CH_4$ (b), and soil $N_2O$ (c) fluxes. The black circle in the center (0,0) of a plot represents the values derived from automated measurements ($F_A$). Blue circles represent estimates from fixed temporal stratification, and red circles represent estimates from an optimized sampling approach (*tuLHs*). Estimates were calculated based on the 258 available automated measurements (Fig. 2), and numeric estimates are in Table S3.

**References**

Bahn, M., Reichstein, M., Davidson, E. A., Gruenzweig, J., Jung, M., Carbone, M. S., Epron, D., Misson, L., Nouvellon, Y., Roupsard, O., Savage, K., Trumbore, S. E., Gimeno, C., Curiel Yuste, J., Tang, J., Vargas, R., and Janssens, I. A.: Soil respiration at mean annual temperature predicts annual total across vegetation types and biomes, Biogeosciences, 7, 2147–2157, 2010.

Ball, B. C.: Soil structure and greenhouse gas emissions: a synthesis of 20 years of experimentation, Eur. J. Soil Sci., 64, 357–373, 2013.

Barba, J., Poyatos, R., and Vargas, R.: Automated measurements of greenhouse gases fluxes from tree stems and soils: magnitudes, patterns and drivers, Sci. Rep., 9, 4005, 2019.

Barba, J., Poyatos, R., Capooci, M., and Vargas, R.: Spatiotemporal variability and origin of $CO_2$ and $CH_4$ tree stem fluxes in an upland forest, Glob. Chang. Biol., 27, 4879–4893, 2021.

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., Crill, P., Curtis, P. S., Davidson, E. A., Desai, A. R., Drake, J. E., El-Madany, T. S., Gavazzi, M., Görres, C., Gough, C. M., Goulden, M., Gregg, J., Gutiérrez del Arroyo, O., He, J., Hirano, T., Hopple, A., Hughes, H., Järveoja, J., Jassal, R., Jian, J., Kan, H., Kaye, J., Kominami, Y., Liang, N., Lipson, D., Macdonald, C. A., Maseyk, K., Mathes, K., Mauritz, M., Mayes, M. A., McNulty, S., Miao, G., Migliavacca, M., Miller, S., Miniat, C. F., Nietz, J. G., Nilsson, M. B., Noormets, A., Norouzi, H., O'Connell, C. S., Osborne, B., Oyonarte, C., Pang, Z., Peichl, M., Pendall, E., Perez-Quezada, J. F., Phillips, C. L., Phillips, R. P., Raich, J. W., Renchon, A. A., Ruehr, N. K., Sánchez-Cañete, E. P., Saunders, M., Savage, K. E., Schrumpf, M.,

585  Scott, R. L., Seibt, U., Silver, W. L., Sun, W., Szutu, D., Takagi, K., Takagi, M., Teramoto,

586  M., Tjoelker, M. G., Trumbore, S., Ueyama, M., Vargas, R., Varner, R. K., Verfaillie, J.,

587  Vogel, C., Wang, J., Winston, G., Wood, T. E., Wu, J., Wutzler, T., Zeng, J., Zha, T., Zhang,

588  Q., and Zou, J.: COSORE: A community database for continuous soil respiration and other

589  soil-atmosphere greenhouse gas flux data, Glob. Chang. Biol., 249, 434, 2020.

590  Bossio, D. A., Cook-Patton, S. C., Ellis, P. W., Fargione, J., Sanderman, J., Smith, P., Wood,

591  S., Zomer, R. J., von Unger, M., Emmer, I. M., and Griscom, B. W.: The role of soil carbon

592  in natural climate solutions, Nature Sustainability, 3, 391–398, 2020.

593  Bowden, R. D., Castro, M. S., Melillo, J. M., Steudler, P. A., and Aber, J. D.: Fluxes of

594  greenhouse gases between soils and the atmosphere in a temperate forest following a

595  simulated hurricane blowdown, Biogeochemistry, 21, 61–71, 1993.

596  Bowden, R. D., Newkirk, K. M., and Rullo, G. M.: Carbon dioxide and methane fluxes by a

597  forest soil under laboratory-controlled moisture and temperature conditions, Soil Biol.

598  Biochem., 30, 1591–1597, 1998.

599  Bréchet, L. M., Daniel, W., Stahl, C., Burban, B., Goret, J.-Y., Salomón, R. L., and Janssens,

600  I. A.: Simultaneous tree stem and soil greenhouse gas ($CO_2$, $CH_4$, $N_2O$) flux

601  measurements: a novel design for continuous monitoring towards improving flux estimates

602  and temporal resolution, New Phytol., 230, 2487–2500, 2021.

603  Butterbach-Bahl, K., Kock, M., Willibald, G., Hewett, B., Buhagiar, S., Papen, H., and Kiese,

604  R.: Temporal variations of fluxes of NO, NO2, N2O, CO2, and CH4 in a tropical rain forest

605  ecosystem, Global Biogeochem. Cycles, 18, https://doi.org/10.1029/2004gb002243, 2004.

606     Capooci, M. and Vargas, R.: Diel and seasonal patterns of soil CO2 efflux in a temperate

607     tidal marsh, Sci. Total Environ., 802, 149715, 2022a.

608     Capooci, M. and Vargas, R.: Trace gas fluxes from tidal salt marsh soils: implications for

609     carbon--sulfur biogeochemistry, Biogeosciences, 19, 4655–4670, 2022b.

610     Capooci, M., Barba, J., Seyfferth, A. L., and Vargas, R.: Experimental influence of storm-

611     surge salinity on soil greenhouse gas emissions from a tidal salt marsh, Sci. Total Environ.,

612     686, 1164–1172, 2019.

613     Castro, M. S., Melillo, J. M., Steudler, P. A., and Chapman, J. W.: Soil moisture as a

614     predictor of methane uptake by temperate forest soils, Can. J. For. Res., 24, 1805–1810,

615     1994.

616     Castro, M. S., Steudler, P. A., Melillo, J. M., Aber, J. D., and Bowden, R. D.: Factors

617     controlling atmospheric methane consumption by temperate forest soils, Global Biogeochem.

618     Cycles, 9, 1–10, 1995.

619     Chilès, J.-P. and Delfiner, P.: Geostatistics: Modeling Spatial Uncertainty, John Wiley &

620     Sons, 720 pp., 2009.

621     Cueva, A., Bullock, S. H., López-Reyes, E., and Vargas, R.: Potential bias of daily soil CO2

622     efflux estimates due to sampling time, Sci. Rep., 7, 11925, 2017.

623     Freeman, C., Lock, M. A., and Reynolds, B.: Fluxes of CO2, CH4 and N2O from a Welsh

624     peatland following simulation of water table draw-down: Potential feedback to climatic

625     change, Biogeochemistry, 19, https://doi.org/10.1007/bf00000574, 1993.

626     Griscom, B. W., Adams, J., Ellis, P. W., Houghton, R. A., Lomax, G., Miteva, D. A.,

627     Schlesinger, W. H., Shoch, D., Siikamäki, J. V., Smith, P., Woodbury, P., Zganjar, C.,

628 Blackman, A., Campari, J., Conant, R. T., Delgado, C., Elias, P., Gopalakrishna, T., Hamsik,

629 M. R., Herrero, M., Kiesecker, J., Landis, E., Laestadius, L., Leavitt, S. M., Minnemeyer, S.,

630 Polasky, S., Potapov, P., Putz, F. E., Sanderman, J., Silvius, M., Wollenberg, E., and

631 Fargione, J.: Natural climate solutions, Proc. Natl. Acad. Sci. U. S. A., 114, 11645–11650,

632 2017.

633 Gunawardana, A., Meek, C., and Xu, P.: A model for temporal dependencies in event

634 streams, Adv. Neural Inf. Process. Syst., 24, 2011.

635 Hao, W. M., Scharffe, D., Crutzen, P. J., and Sanhueza, E.: Production of N2O, CH4, and

636 CO2 from soils in the tropical savanna during the dry season, J. Atmos. Chem., 7, 93–105,

637 1988.

638 Hill, A. C., Barba, J., Hom, J., and Vargas, R.: Patterns and drivers of multi-annual CO2

639 emissions within a temperate suburban neighborhood, Biogeochemistry, 152, 35–50, 2021.

640 Huntington, D. E. and Lyrintzis, C. S.: Improvements to and limitations of Latin hypercube

641 sampling, Probab. Eng. Mech., 13, 245–253, 1998.

642 Hutchinson, G. E.: The Concept of Pattern in Ecology, Proceedings of the Academy of

643 Natural Sciences of Philadelphia, 105, 1–12, 1953.

644 Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., Kholod, N.,

645 Manzon, J., Marchesi, R., Paredes, D., and Bond-Lamberty, B.: A restructured and updated

646 global soil respiration database (SRDB-V5), Data, Algorithms, and Models,

647 https://doi.org/10.5194/essd-2020-136, 2020.

648 Keller, M., Kaplan, W. A., and Wofsy, S. C.: Emissions of N2O, CH4 and CO2 from tropical

649 forest soils, J. Geophys. Res., 91, 11791, 1986.

650 Kim, D. G., Vargas, R., Bond-Lamberty, B., and Turetsky, M. R.: Effects of soil rewetting

651 and thawing on soil gas fluxes: a review of current literature and suggestions for future

652 research, Biogeosciences, 9, 2459–2483, 2012.

653 Le, V. H., Díaz-Viera, M. A., Vázquez-Ramírez, D., del Valle-García, R., Erdely, A., and

654 Grana, D.: Bernstein copula-based spatial cosimulation for petrophysical property prediction

655 conditioned to elastic attributes, J. Pet. Sci. Eng., 193, 107382, 2020.

656 Lucas-Moffat, A. M., Huth, V., Augustin, J., Brümmer, C., Herbst, M., and Kutsch, W. L.:

657 Towards pairing plot and field scale measurements in managed ecosystems: Using eddy

658 covariance to cross-validate CO2 fluxes modeled from manual chamber campaigns, Agric.

659 For. Meteorol., 256–257, 362–378, 2018.

660 Luo, G. J., Kiese, R., Wolf, B., and Butterbach-Bahl, K.: Effects of soil temperature and

661 moisture on methane uptake and nitrous oxide emissions across three different ecosystem

662 types, Biogeosciences, 10, 3205–3219, 2013.

663 Murguia-Flores, F., Arndt, S., Ganesan, A. L., Murray-Tortarolo, G., and Hornibrook, E. R.

664 C.: Soil Methanotrophy Model (MeMo v1.0): a process-based model to quantify global

665 uptake of atmospheric methane by soil, Geosci. Model Dev., 11, 2009–2032, 2018.

666 Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., and Erasmi, S.: Greenhouse gas

667 emissions from soils—A review, Geochem. Explor. Environ. Analy., 76, 327–352, 2016.

668 Ojanen, P., Minkkinen, K., Alm, J., and Penttilä, T.: Soil–atmosphere CO2, CH4 and N2O

669 fluxes in boreal forestry-drained peatlands, For. Ecol. Manage., 260, 411–421, 2010.

670  Petrakis, S., Seyfferth, A., Kan, J., Inamdar, S., and Vargas, R.: Influence of experimental

671  extreme water pulses on greenhouse gas emissions from soils, Biogeochemistry, 133, 147–

672  164, 2017.

673  Petrakis, S., Barba, J., Bond-Lamberty, B., and Vargas, R.: Using greenhouse gas fluxes to

674  define soil functional types, Plant Soil, 423, 285–294, 2018.

675  Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J. W., Todd-

676  Brown, K., and Vargas, R.: The value of soil respiration measurements for interpreting and

677  modeling terrestrial carbon cycling, Plant Soil, 413, 1–25, 2017.

678  Pyrcz, M. J. and Deutsch, C. V.: Geostatistical Reservoir Modeling, OUP USA, 433 pp.,

679  2014.

680  R Code Team: R: A language and environment for statistical computing, 2013.

681  Rowlings, D. W., Grace, P. R., Kiese, R., and Weier, K. L.: Environmental factors

682  controlling temporal and spatial variability in the soil-atmosphere exchange of $CO_2$, $CH_4$

683  and $N_2O$ from an Australian subtropical rainforest, Glob. Chang. Biol., 18, 726–738, 2012.

684  Savage, K., Phillips, R., and Davidson, E.: High temporal frequency measurements of

685  greenhouse gas emissions from soils, Biogeosciences, 11, 2709–2720, 2014.

686  Shakoor, A., Shahbaz, M., Farooq, T. H., Sahar, N. E., Shahzad, S. M., Altaf, M. M., and

687  Ashraf, M.: A global meta-analysis of greenhouse gases emission and crop yield under no-

688  tillage as compared to conventional tillage, Sci. Total Environ., 750, 142299, 2021.

689  Storn, R. and Price, K.: Differential Evolution – A Simple and Efficient Heuristic for global

690  Optimization over Continuous Spaces, J. Global Optimiz., 11, 341–359, 1997.

691 Tallec, T., Brut, A., Joly, L., Dumelié, N., Serça, D., Mordelet, P., Claverie, N., Legain, D.,

692 Barrié, J., Decarpenterie, T., Cousin, J., Zawilski, B., Ceschia, E., Guérin, F., and Le Dantec,

693 V.: N2O flux measurements over an irrigated maize crop: A comparison of three methods,

694 Agric. For. Meteorol., 264, 56–72, 2019.

695 Tang, X., Liu, S., Zhou, G., Zhang, D., and Zhou, C.: Soil-atmospheric exchange of CO2 ,

696 CH4 , and N2 O in three subtropical forest ecosystems in southern China, Glob. Chang. Biol.,

697 12, 546–560, 2006.

698 Trangmar, B. B., Yost, R. S., and Uehara, G.: Application of Geostatistics to Spatial Studies

699 of Soil Properties, in: Advances in Agronomy, vol. 38, edited by: Brady, N. C., Academic

700 Press, 45–94, 1986.

701 Ullah, S. and Moore, T. R.: Biogeochemical controls on methane, nitrous oxide, and carbon

702 dioxide fluxes from deciduous forest soils in eastern Canada, J. Geophys. Res., 116,

703 https://doi.org/10.1029/2010jg001525, 2011.

704 Vargas, R.: How a hurricane disturbance influences extreme CO2 fluxes and variance in a

705 tropical forest, Environ. Res. Lett., 2012.

706 Vargas, R., Detto, M., Baldocchi, D. D., and Allen, M. F.: Multiscale analysis of temporal

707 variability of soil CO2 production as influenced by weather and vegetation, Glob. Chang.

708 Biol., 16, 1589–1605, 2010.

709 Vargas, R., Carbone, M. S., Reichstein, M., and Baldocchi, D. D.: Frontiers and challenges in

710 soil respiration research: from measurements to model-data integration, Biogeochemistry,

711 102, 1–13, 2011.

712    Vargas, R., Sánchez-Cañete P., E., Serrano-Ortiz, P., Curiel Yuste, J., Domingo, F., López-

713    Ballesteros, A., and Oyonarte, C.: Hot-Moments of Soil CO2 Efflux in a Water-Limited

714    Grassland, Soil Systems, 2, 47, 2018.

715    Vicca, S., Bahn, M., Estiarte, M., van Loon, E. E., Vargas, R., Alberti, G., Ambus, P., Arain,

716    M. A., Beier, C., Bentley, L. P., Borken, W., Buchmann, N., Collins, S. L., de Dato, G.,

717    Dukes, J. S., Escolar, C., Fay, P., Guidolotti, G., Hanson, P. J., Kahmen, A., Kröel-Dulay, G.,

718    Ladreiter-Knauss, T., Larsen, K. S., Lellei-Kovacs, E., Lebrija-Trejos, E., Maestre, F. T.,

719    Marhan, S., Marshall, M., Meir, P., Miao, Y., Muhr, J., Niklaus, P. A., Ogaya, R., Peñuelas,

720    J., Poll, C., Rustad, L. E., Savage, K., Schindlbacher, A., Schmidt, I. K., Smith, A. R., Sotta,

721    E. D., Suseela, V., Tietema, A., van Gestel, N., van Straaten, O., Wan, S., Weber, U., and

722    Janssens, I. A.: Can current moisture responses predict soil CO2 efflux under altered

723    precipitation regimes? A synthesis of manipulation experiments, Biogeosciences, 11, 2991–

724    3013, 2014.

725    Villarreal, S., Guevara, M., Alcaraz-Segura, D., and Vargas, R.: Optimizing an

726    Environmental Observatory Network Design Using Publicly Available Data, J. Geophys.

727    Res. Biogeosci., 124, 1812–1826, 2019.

728    Wang, G. and Chen, S.: A review on parameterization and uncertainty in modeling

729    greenhouse gas emissions from soil, Geoderma, 170, 206–216, 2012.

730    Warner, D. L., Guevara, M., Inamdar, S., and Vargas, R.: Upscaling soil-atmosphere CO2

731    and CH4 fluxes across a topographically complex forested landscape, Agricultural and forest,

732    264, 80–91, 2019.

733    Werner, C., Kiese, R., and Butterbach-Bahl, K.: Soil-atmosphere exchange of N2O, CH4,

734    and CO2and controlling environmental factors for tropical rain forest sites in western Kenya,

735    J. Geophys. Res., 112, https://doi.org/10.1029/2006jd007388, 2007.

736    Wu, X., Brüggemann, N., Gasche, R., Shen, Z., Wolf, B., and Butterbach-Bahl, K.:

737    Environmental controls over soil-atmosphere exchange of N2O, NO, and CO2in a temperate

738    Norway spruce forest, Global Biogeochem. Cycles, 24,

739    https://doi.org/10.1029/2009gb003616, 2010.

740    Yao, Z., Zheng, X., Xie, B., Liu, C., Mei, B., Dong, H., Butterbach-Bahl, K., and Zhu, J.:

741    Comparison of manual and automated chambers for field measurements of N2O, CH4, CO2

742    fluxes from cultivated land, Atmos. Environ., 43, 1888–1896, 2009.