Monte Carlo Simulations to Assess the Uncertainty of Locating and Quantifying CO2 Leakage Flux from Deep Geological or Anthropogenic Sources.

Stan E. Beaubien[a], Giancarlo Ciotoli[b], Maria Grazia Finoia[c], Salvatore Lombardi[a], Sabina Bigi[a]

[a]Dipartimento di Scienze della Terra, Università di Roma “La Sapienza”, Piazzale Aldo Moro 5, Rome 00185, Italy.

[b]Istituto di Geologia Ambientale e Geoingegneria del Consiglio Nazionale delle Ricerche, Strada Provinciale 35d, Montelibretti 00010, Italy.

[c]Istituto Superiore per la Protezione e la Ricerca Ambientale, Via Vitaliano Brancati 48, Rome 00144, Italy.

Author	email	ORCID
Stan E. Beaubien	stanley.beaubien@uniroma1.it	0000-0001-6127-9744
Giancarlo Ciotoli	giancarlo.ciotoli@igag.cnr.it	0000-0002-9496-3021
Maria Grazia Finoia	mariagrazia.finoia@isprambiente.it	0000-0003-0765-0631
Salvatore Lombardi	salvatore.lombardi@gmail.com	
Sabina Bigi	sabina.bigi@uniroma1.it	0000-0003-1575-2689

Corresponding author: Sabina Bigi (sabina.bigi@uniroma1.it)

Acknowledgments: The assistance of Pietro Sacco and Davide de Angelis for the collection of the Latera field data is gratefully acknowledged.

Funding: The research leading to these results received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653718.

Conflicts of interest/Competing interests: The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material: Datasets and configuration files used to conduct the reported simulations are available from the Zenodo open-source online repository at https://doi.org/10.5281/zenodo.4573725 (Beaubien and Bigi, 2021b).

Code availability: The MC-Flux (V1.0) installation package and user’s manual are available from the Zenodo open-source online repository at https://doi.org/10.5281/zenodo.4573575 (Beaubien and Bigi, 2021a).

Authors' contributions: Conceptualization [SEB, SL], Methodology, Software, Visualization and Writing – Original Draft [SEB]; Formal Analysis [SEB, GC, MGF]; Validation [GC, MGF]; Writing – Review and Editing [GC, MGF, SB]; Supervision [SL]; Project administration and Funding acquisition [SB, SL].

Abstract

Accurately locating and quantifying CO_2 leakage to the atmosphere is important for diffuse degassing studies in volcanic / geothermal areas and for safety monitoring of Carbon Capture and Storage (CCS) sites. This is typically conducted by measuring CO_2 flux at numerous points over a large area and applying statistics or geostatistical interpolation. Probability and accuracy of the results will depend on many factors related to survey/data-processing choices and site characteristics, and thus uncertainties can be difficult to quantify. To address this issue, we have developed a Monte Carlo-based program (MC-Flux) that repeatedly subsamples a high-resolution synthetic or real dataset using five different sampling strategies at multiple user-defined sample densities, keeping track of the anomalies found and estimating total flux using four approaches from the literature. This paper describes the use of MC-Flux to assess the potential impact of various sampling and interpretation decisions on the accuracy of the final results. Simulations show that an offset grid sample distribution yields the best results, however relatively dense sampling is required to obtain a high probability of an accurate flux estimate. For the test dataset used, ordinary kriging interpolation produces a range of flux estimates that are centered on the true value while sequential Gaussian simulation tends to slightly overestimate values at intermediate sample spacings and is sensitive to input parameters. These results point to the need for developing new approaches that decrease uncertainty, such as integration with high-resolution co-kriging datasets that complement the more accurate point flux measurements.

Keywords: Soil diffuse degassing, Carbon Capture and Storage (CCS), CO_2 flux quantification, uncertainty, probability, Monte Carlo

1. Introduction

Accurately finding and quantifying the leakage of deep-origin carbon dioxide (CO_2) to the atmosphere is challenging considering that point flux measurements conducted over large areas are used to characterize spatial anomalies that can be irregular in size, shape, magnitude and distribution and that are superimposed on a background biological flux that is both spatially and temporally variable. In this paper we use a Monte Carlo approach to study the impact of these variables and to quantify some of the related uncertainties.

Initial efforts to determine the probability of finding an anomaly using gridded sampling were conducted in the field of mineral exploration. Singer (1972, 1975) used a geometric probability model called ElipGrid to show how triangular and square grids are often equivalent, but that the former can be up to 6% more efficient for certain conditions. The accuracy of this code was proven using a Monte Carlo approach (Davidson 1995), followed by its integration into an environmental software package called Visual Sample Plan (VSP) (Matzke et al. 2014). Instead, the probability of finding anomalies using random sampling, calculated using the formula:

$$P=1-\left[1-x/A\right]^n, \quad \textbf{(1)}$$

where x is the size of the anomaly and A is the size of the survey area, is low and the approach is generally inefficient (Oldenburg et al. 2003).

Uncertainties in quantifying CO_2 leakage can be subdivided into four categories. First, analytical uncertainty is related to site conditions (Bain et al. 2005) and sampling / data processing protocols (Elío et al 2012; Kutzbach et al. 2007). Estimates of the measurement reproducibility range from ±10% (Chiodini et al. 1998) to ±24% for low flux rates (Carapezza and Granieri 2004). Second, the probability of finding anomalies depends on their size and shape versus sample density and distribution. Wong (2018) used a Monte Carlo approach to illustrate that a large number of random samples are needed to obtain a high probability of accurately estimating the average CH_4 flux above a small landfill site.

Third, converting the measured point flux values into an estimated total flux for the entire survey area requires statistical characterization (e.g., arithmetic mean, AM, or minimum variance unbiased estimator, MVUE) or geostatistical extrapolation (e.g., kriging or sequential Gaussian simulation, SGS). Lewicki et al. (2005) compared these four approaches and found them to yield similar results but recommended SGS because this stochastic method honors the histogram and variogram of the original data and provides a measure of the uncertainty. Elío et al. (2016) also found SGS to yield the most robust estimates if the data can be fit with a variogram, otherwise they recommend AM, MVUE or bootstrap resampling for normal, log-normal or mixed populations, respectively. Cardellini et al. (2003) used SGS and data from multiple sites to define an empirical relationship linking flux estimate precision with the number of samples needed within a circular area having a radius equal to the data's variogram range. Schroder et al. (2017) found that AM and MVUE approaches did not accurately estimate leakage rates from a controlled release experiment and thus proposed a new cubic interpolation approach that is not limited by spatial or statistical distribution . At present, SGS is the most commonly used approach, although it too has its limitations (Caers 2000; Gilardi et al. 2002; Emery 2004; Paravarzar et al. 2015).

Fourth, the background biological flux must be estimated and subtracted from the total flux to determine the leakage flux. This can been conducted by making measurements in a separate non-leaking area (Chiodini et al. 2007), but it is much more common to use log-normal probability plots to separate populations of different origins (Chiodini et al. 2020). Elío et al. (2016), however, point out that this approach can be subjective and instead recommend a maximum likelihood procedure. It is extremely difficult to estimate the background uncertainty, given its spatial and temporal variability (Sainju et al. 2008; Leon et al. 2014; Bond-Lamberty et al. 2019), however it can have an important impact on the final leakage uncertainty if background flux represents a significant proportion of the total. Recent efforts to use components associated with the leaking CO_2, such as isotopic signatures or radon (e.g., Bini et al. 2019; Viveiros et al. 2020), have shown how this uncertainty can be significantly reduced. If resources are limited, however, additional analyses may not be feasible or they could lead to a reduction in the total number of flux measurement points (thus reducing spatial resolution).

The present work takes a unique approach to quantitatively assess the issues and approaches described above. It better develops and greatly extends the preliminary efforts by Beaubien et al. (2009) and Beaubien (2015) to develop a Monte Carlo – based code (Beaubien and Bigi 2021a) capable of defining the probability of locating leakage flux anomalies and assessing the accuracy of associated quantification. Synthetic and real data (Beaubien and Bigi 2021b) are used to explore the effectiveness of different sampling strategies, the impact of anomaly shape and orientation, and to assess the uncertainties linked to the multiple steps (sampling density and strategy decisions, interpolation, background subtraction) required to estimate a final leakage flux. Ideas are presented that could reduce these uncertainties without significantly increasing survey times or costs.

2. Materials and Methods

The MC-Flux program, coded in Visual Basic 6 and run in Windows 10, uses a Monte Carlo approach to study the influence of sampling strategies and sample spacing on the probability of finding a gas leakage anomaly and on the accuracy of leakage flux estimates.

Three main functions are available. The first imports or creates high-resolution input data, sub-samples it a user-defined number of times at different spacings using one of five sampling strategies (keeping track of anomalies or calculating statistics to estimate total flux), and saves each sub-sampled equi-probable realization in a simple text file. This function can be used as a stand-alone for determining the probabilities of finding leakage anomalies, or the total flux can be estimated for each sub-sampled realization both statistically and geostatistically using external programs. The second function calls the commercial software Surfer (Golden Software) to perform interpolation using the ordinary kriging, natural neighbor, inverse distance to a power or the radial basis function methods, although only ordinary kriging was applied here. Note that the user must have Surfer installed to access this functionality; development was performed using version 9 however limited testing has shown that MC-Flux also works with version 20. The third function calls the program sGsim from the open source GSLib library (Deutsch and Journel 1997) to conduct sequential Gaussian simulations; this library of DOS executable files can be freely downloaded (http://www.gslib.com) and copied onto the user's computer.

A flowchart showing the logical structure of the program (Figure S1), the graphical user interface (GUI) developed for selecting options and inputting parameters (Figure S2), and the MC-Flux user manual are given in the Supporting Materials. Note that all GUI input is saved to a configuration file that can be re-loaded to repeat a simulation or make systematic parameter changes.

2.1. Data input

Both synthetic and real data can be used as input for estimating total CO2 flux while only synthetic data can be used to calculate the probabilities of finding anomalies.

Synthetic data is created in two steps. First the background flux field is generated using one of three options: i) all points are assigned a value of 0 (to facilitate recognizing individual anomalies); ii) an internally created normal distribution is randomly placed over the grid (Figure S3a); or iii) an external file is imported into the program. For this last option, we have created a log-normal distribution using Poptools (Hood 2010) and coherently distributed it over a 1km^2 grid assuming an inverse relationship between topography and biological CO_2 flux (Figure S3b, S4), given that topography influences environmental parameters (e.g., water content, organic matter) that affect soil respiration (Riveros-Iregui and McGlynn 2009). Second, one or multiple, chosen or random, circular or elliptical, gas leakage areas ("vents") are superimposed on the background based on user-defined parameters (number of vents, location, semi-major axis length, semi-major vs semi-minor axis ratio, orientation, and maximum flux rate). This information is saved and can be imported for subsequent simulations. For the "find anomaly" option, vent points are assigned the vent number in a background of zeros, which allows for individual anomalies to be recognized. For the "calculate flux" option, individual vent flux values are calculated using an empirically defined formula based on profiles across multiple gas vents (Figure S5a,b) and plotted on a background flux field (Figure S5c, d).

If real data is used, the program imports a text file consisting of a single column of values ordered sequentially for each X point along progressive Y lines. As it is not feasible to collect real data at 1m sample spacing over a large area, a dataset of 548 CO_2 flux measurements made on a regular, 10 m spacing grid in a polygonal area was modified for this purpose. The survey area is located in a well-studied field containing multiple gas vents in the Latera Caldera, Italy (Beaubien et al. 2008; Pettinelli et al. 2010). Flux measurements were made on July 13th and 14th, 2014, using three in-house accumulation chamber units that were inter-calibrated prior to the survey. The original grid was extended to a 260 x 375 m area by assigning random background values outside the polygon, and then a high-resolution (HR) grid (1 m node spacing) was created using the average of 10 sequential Gaussian simulations (SGS; Deutsch and Journel 1997); SGS was performed to minimize any spatial biasing that may result from gridded sampling (Figure S6a). This data is formed by two main populations (Figure S6b). The lower background population ranges from about 20 to 70 g m^{-2} d^{-1} and has a mean of 51 g m^{-2} d^{-1}; these relatively high biogenic fluxes were related to rainfall events shortly before sampling, similar to that observed by Viveiros et al. (2020). This interpretation is supported by the much lower background values observed in the same field (Figure S6c) during a previous campaign in July, 2006 (Annunziatellis et al. 2008). The upper, leakage-related population has a maximum value of 1645 g m^{-2} d^{-1}. The standardized variogram of the HR data consists of a spherical model with a nugget of 0.02, a variance of 0.98 and a range of 37m (Figure S6d).

2.2. Sub-sampling

The program subsamples the high-resolution flux array using one of five possible sampling strategies (Figure S7): i) square grid, where X and Y distances are the same and points are aligned orthogonally; ii) off-set grid, where X and Y distances are the same but every second row is offset horizontally by X/2; iii) triangular grid, where every second row is offset horizontally by X/2 but with all point-to-point distances the same, resulting in $Y<X$; iv) random grid, where the program steps through a square grid, but at each node a random point is selected within a user-defined radius; and v) purely random sampling. Each subsampling iteration produces a single, equi-probable, sub-sample dataset.

Simulations are conducted for different "sample densities", which are defined by the sample spacing (grids) or equivalent number of samples (random) chosen by the user. A nested loop structure produces N unique sub-sampling realizations for M sample densities. Note, however, that the number of unique subsample realizations N for the non-random grid methods is limited to (sample spacing)^2 when the input data has a 1m spacing, meaning that fewer unique simulations can be performed for closely spaced grids.

2.3. Calculations

For the "find anomaly" simulations, each sub-sampled dataset is queried to determine if any sample points intersect an anomaly, and if yes which ones. After all realizations for that sampling density are completed, the probability of finding each individual anomaly and the average number of anomalies found is calculated.

For the "calculate flux" simulations, the first step involves calculating the "true" leakage flux by subtracting the background flux from the total flux for the original HR dataset; this value is then used as a benchmark for the subsequent subsampling results. For synthetic data the background and background-plus-vents flux rates are calculated using:

$$\boldsymbol{\varphi_T} = \textstyle\sum_1^n (\boldsymbol{A_c} * \boldsymbol{\varphi_c}) \quad \textbf{(2)}$$

Where φ_T is the total flux being calculated, n is the total number of grid cells, A_c is the surface area of each cell (m^2), and φ_c is the CO_2 flux for each cell (g m^{-2} d^{-1}). For real data, the total flux is calculated using Eqn 2 while the background is calculated by multiplying the grid's surface area by the average of the lower flux population in a log-normal probability plot of the entire raw dataset (Chiodini et al. 2007).

Four different approaches are used to calculate the total flux of the sub-sampled realizations, two statistical (arithmetic mean, AM, and minimum variance unbiased estimator, MVUE) and two geostatistical (ordinary kriging, OK, and sequential Gaussian simulations, SGS). For both statistical methods the resultant value is multiplied by the total surface area, and then the previously calculated "true" background leakage is subtracted from this value to yield the total leakage flux estimate. The AM, which is most appropriate for normally distributed datasets, is calculated as $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. The MVUE, $\hat{\mu}_1$, which better describes log-normal distributions, is calculated as $\hat{\mu}_1 = [\exp(\overline{y})]\, \psi_n(t)$ (Gilbert 1987; Elío et al. 2016), where $\overline{y}$ is the sample arithmetic mean calculated using the n log-transformed values, t is half of the log transformed variance s_y^2, and ψ_n is an infinite series (see Figure S8 for details).

Regarding the geostatistical methods, OK is implemented using the commercial software package Surfer 9 (Golden Software) and SGS is implemented using the program sGsim within the open source GSLib library (Deutsch and Journel 1997). The OK approach yields a single array for the grid area, whereas SGS produces a user-defined number of realizations for each sub-sampled dataset that are then averaged to produce a single array. As a compromise between level of detail and processing time, an output cell size of 5 x 5 m was chosen for all simulations. Each individual array produced by both OK and SGS is then processed with Eqn 2 to calculate the total flux, and the previously calculated "true" background flux is subtracted to yield the associated leakage flux estimate for each realization.

3. Results and Discussion

3.1. Probability of finding an anomaly

3.1.1. Comparison with literature

The accuracy of the MC-Flux Monte Carlo probability estimates were validated using test data reported in Table 1 in Singer (1972) and Table A3 in Davidson (1995) that consist of a wide range of single ellipse sizes, shapes, and orientations coupled with different node spacings for square and triangular grid strategies. The MC-Flux results, based on 5000 realizations, have a 1:1 linear correlation with those generated using the mathematical approach of the ElipGrid program, with $R^2 = 0.9999$ for both datasets and most individual simulations within 0.7% (Figure S9a). The test data were modified to convert grid spacing into an equivalent total number of samples to test the purely random sampling approach. The MC-Flux results match very closely with the theoretical probability calculated using Eqn 1 (Figure S9b), with $R^2 = 0.9998$ and most individual simulations within 1%. These results give confidence in the MC-Flux simulations reported below.

3.1.2. Number of simulations necessary

Tests were performed to determine how many simulations are needed to obtain an stable probability estimate. All five strategies were used to sub-sample two different 1 km^2 grids, one containing a circular anomaly (a=56.42, x/A=0.01) that was sampled with a c. 120m spacing and one with an elliptical anomaly (a=79.79, b/a=0.5, angle=22.5°, x/A=0.01) that was sampled with a c. 150m spacing. Each test involved 10,000 realizations, each of which yielded 1 or 0 if the anomaly was encountered or not. A total of "n" random samples were collected from this dataset (where n = 30, 50, 75, 100, 200, 300, 400, 500, 1000, 2000, or 3000) a total of 200 times each to calculate the probability of finding the anomaly using different numbers of simulations (i.e., "n").

All square grid results for the first test yield relatively similar average probabilities but, as expected, the total range of estimated probability narrows significantly with the inclusion of larger numbers of simulations (Figure S10a). The standard deviation of all tests show a linear decrease with increasing number of simulations on a log-log plot (Figure S10b), regardless of anomaly shape, sampling strategy or sample spacing. Based on these results and the fact that the sub-sampling algorithm is fast, all "find anomaly" probability tests reported below were conducted using 5000 simulations ($1\sigma \approx 0.7\%$). In contrast, the OK and SGS flux estimation tests are much slower, and thus 100 ($1\sigma \approx 5\%$) or 500 ($1\sigma \approx 2\%$) simulations were used for these.

3.1.3. Probability of finding a single circular anomaly

The efficiency of the different sampling strategies was first tested for the simple case of a single circular anomaly (a=56.42 m; x=10,000 m^2) in the middle of a 1 km^2 area (x/A=0.01). Simulations were conducted so that each strategy used the same average number of samples, thus spacing was the same for the square, offset and random grids (from 40 to 150 m), 1.07457 times larger for the triangular grid (Singer 1975), and an equivalent number of samples was used for the purely random method.

Random sampling gives significantly poorer results across all probability levels (Figure 1). All four gridded methods yield similar results up to a probability of 0.45, with the random grid rapidly diverging at higher levels. Its trend is a function of the chosen random search radius (here 50%), such that low values move its trend closer to the square grid results while higher values move it towards the random results. At a probability level above ca 0.8 the square grid method becomes less efficient than the offset and triangular grids, with the latter two being essentially equivalent over all probability levels. This higher efficiency of the triangular versus square grid is in agreement with the maximum improvement of about 6% observed by Singer (1975). As a comparison, the dashed lines in Figure 1 show that to ensure a 95% probability of finding the anomaly, about 100 samples would need to be collected using offset and triangular grids, 114 using the square grid, 148 using the random grid with a 50% search radius, and 300 using a purely random approach. Sample numbers can also be estimated for other conditions using the ratio of the anomaly and grid cell areas given in the upper X axis.

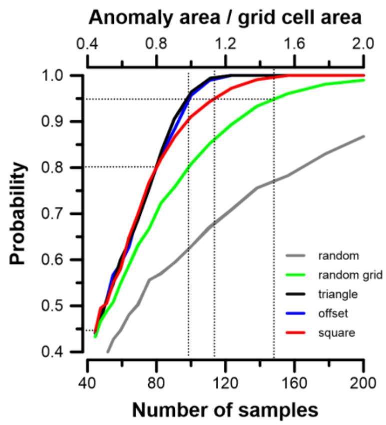

Figure 1. Probability of finding a circular anomaly as a function of sampling strategy and sample spacing. The lower X axis refers to the modelled conditions of x=10,000 and A=1,000,000 while the upper X axis standardizes the trends for all anomaly/cell ratios.

3.1.4. Probability at different ellipse angles

The probability of finding an ellipse as a function of its orientation was tested using the five different sampling strategies (spacing from 40 -175 m in 5 m steps, or equivalent). Because MC-Flux can keep track of individual anomalies and because the location of an anomaly within the domain has no effect on the Monte Carlo results, 46 ellipses (x=10,000 m^2, b/a =0.25) were created at angles from 0-90° at 2° intervals within a 1 km^2 area (Figure S11).

Using the square grid results as an example, orientation and sample spacing clearly have a significant combined effect on the probability of finding a narrow ellipse (Figure 2a). At very low sampling densities the trends are relatively flat, while at very high densities all directions tend towards a probability of 100%. Instead, between these extremes maximum probabilities occur at 28° and 62°, minimum probabilities occur at 0° and 90°, and a lesser minimum occurs at 45°, all due to the geometric positioning of the ellipse within the distributed points. For some sample spacings the differences are very large; for example, probabilities at a 95 m spacing can range from 60% at 0° up to almost 100% at 28°. This could mean that between 100 to 300 samples would be needed to attain 95% probability (Figure 2b) depending on orientation. Similar plots for all sampling strategies are given in Figure S12.

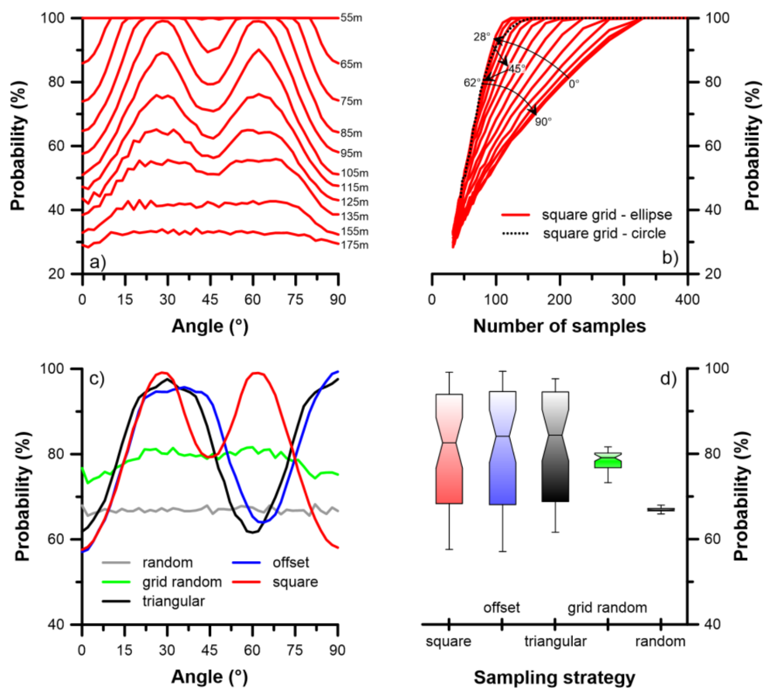

Figure 2. Plots showing the impact of orientation on the probability of finding a 10,000 m^2, elliptical (b/a=0.25) anomaly in a 1 km^2 area: (a) square grids of different spacings (reported numbers); (b) same data plotted against the number of samples, with the dashed line showing the circular anomaly trend from Figure 1; (c) probabilities for all five sampling strategies at an equivalent sampling distance of 95 m; and (d) statistical distribution of the data in (c).

Results from a sample spacing of 95 m (or equivalent) are used to compare the relative trends of all five strategies (Figure 2c). Again the offset and triangular grid results are quite similar (with a slight shift of < 5°), but are quite different from the square grid results. In particular, maximum and minimum values are opposite at 60° and 90°. The random grid results are only weakly influenced by anomaly direction and have probability values on the order of 80%, while random sampling, as expected, is not affected by orientation and has a low, flat trend around 68%. The statistical distribution of these data (Figure 2d) show how the square, triangular, and offset grid strategies are essentially equivalent, meaning that if the angle is unknown or variable none of these methods are superior to the others for finding this narrow anomaly. Instead, the random grid produces a median probability that is only slightly lower than the other grid methods but with a much narrower range (i.e., more predictable outcome). Although potentially advantageous, this approach is not necessarily practical for field sampling campaigns. Similar plots for 75 m and 135 m sample spacing, or equivalent, are given in Figure S13.

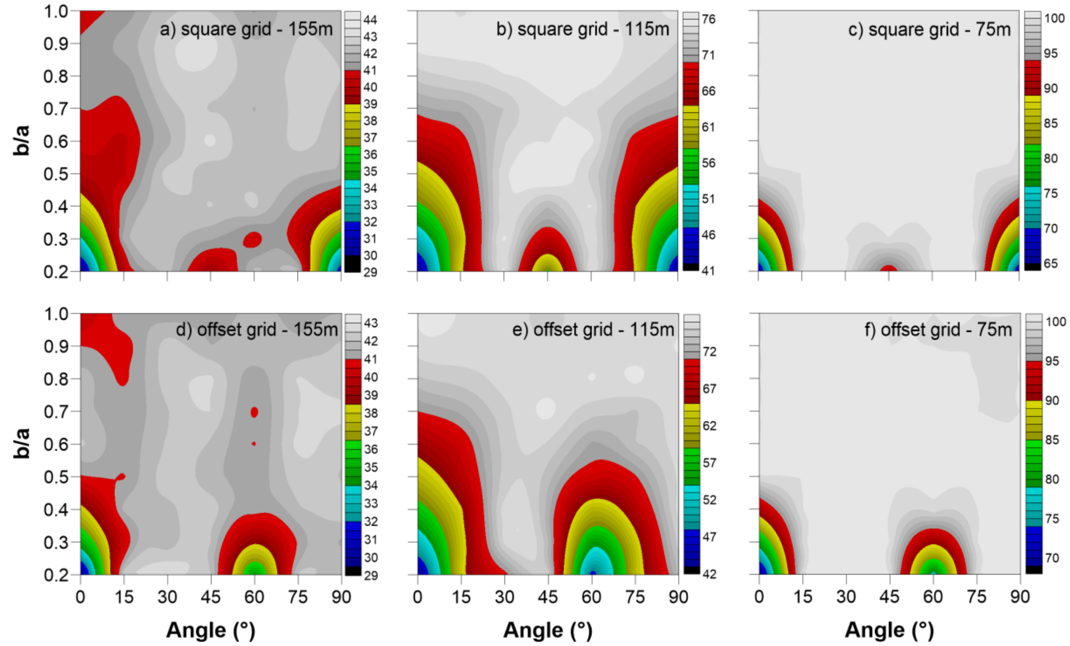

Figure 3. Plots showing the probability of finding a 10,000 m^2 elliptical anomaly in a 1,000,000 m^2 area as a function of its shape (b/a) and angle (where horizontal is 0°, moving counter-clockwise to vertical at 90°) for the labelled grid types and sample spacings.

Further simulations were conducted to address the combined effect of b/a values from 0.2 to 1.0 (0.1 step size) and angles from 0° to 90° (15° step size). A single domain was defined that contained all anomalies (Figure S14) and simulations were conducted with the square and offset grid methods for sample spacings from 40 to 175 m (5 m step size). Trends are similar over a wide range of sample spacings and the impact of anomaly orientation diminishes at higher b/a values (Figure 3). Using the square grid results as an example, the total range of probability values is narrowest for the widest spacing (e.g., 15% difference in Figure 3a) but is much greater for the other two spacings (e.g., 35% in both Figure 3 b and c). A comparison of the latter two also shows that the impact is observed at much greater b/a values for the 115 m grid spacing (up to b/a=0.7 for a 5% difference) than for the 75 m spacing (b/a=0.4), as progressively narrower spacing is more likely to intersect wider ellipses that may lie between rows or columns. The square grid results are highly symmetrical at 0° and 90° (e.g., Figure 3b) whereas the offset grid results show higher probabilities at higher b/a values for angles at 0° compared 60° (e.g., Figure 3e). That said, the average probability for all b/a values and orientation angles for a given sample spacing are essentially the same for the square and offset grids.

3.1.5. Average number of anomalies found

Considering that more than one leakage anomaly may occur in a given survey area, simulations were performed to assess how many can be located using different sample spacings. Initial tests involved 6 different scenarios containing 1, 2, 4, 6, 8, or 10 circular anomalies, with all anomalies in each individual scenario having the same size and the sum of their areas being equal to 10,000 m^2 (i.e., $x_{total}/A=0.01$) (Figure S15). The square and offset grid methods were used for each sample spacing; as results are very similar only the latter results are described here.

When plotted in terms of the average number of anomalies found, all scenarios follow the same trends for both sample spacing (Figure 4a) and total number of samples collected (Figure 4b). Instead, plotting the same data in terms of the percentage of anomalies found relative to the total number present illustrates how, as expected, the smaller the average anomaly size the smaller the proportion that are found (Figure 4c, d). For example, the dashed lines in Figure 4a and c show how an average of 3 anomalies are found in the 4, 6, 8, and 10 anomaly scenarios when a sample spacing of about 60 m is used, but that this represents 70%, 48%, 35%, and 28%, respectively, of the anomalies actually present. Clearly how a total leakage amount is subdivided and, in turn, how many of these anomalies are found, will have an important impact on eventual quantification estimates.

Using the 10 anomaly simulation as the base case, two additional tests were performed to look at the impact of different sized anomalies. Centered on the semi-major axis value from this scenario 1 (i.e., a=17.8 m), scenario 2 anomalies increase from a=12.6 to 22.5 m in ca. 1 m steps while scenario 3 anomalies increase from a=7.4 to 26.1 m in ca. 2 m steps (Figure S16). As above, the total surface area of all 10 anomalies in each scenario equals 10,000 m^2. As the distribution of sizes is symmetrical above and below the scenario 1 dimension, trends for all three scenarios are the same up to 50% of the anomalies found (Figure 4e, f). Deviations are observed above this threshold, however, as the smaller anomalies require a closer grid spacing to ensure their discovery. For example, in order to find 9 out of 10 anomalies, 1000 samples are needed for the base case, 1300 for scenario 2 and 2500 for scenario 3, which corresponds to a grid spacing of 32, 27 and 19 m, respectively. While intuitive, these results quantitatively illustrate how challenging it is to locate all leakage areas present, especially if small anomalies occur in a wider size distribution.

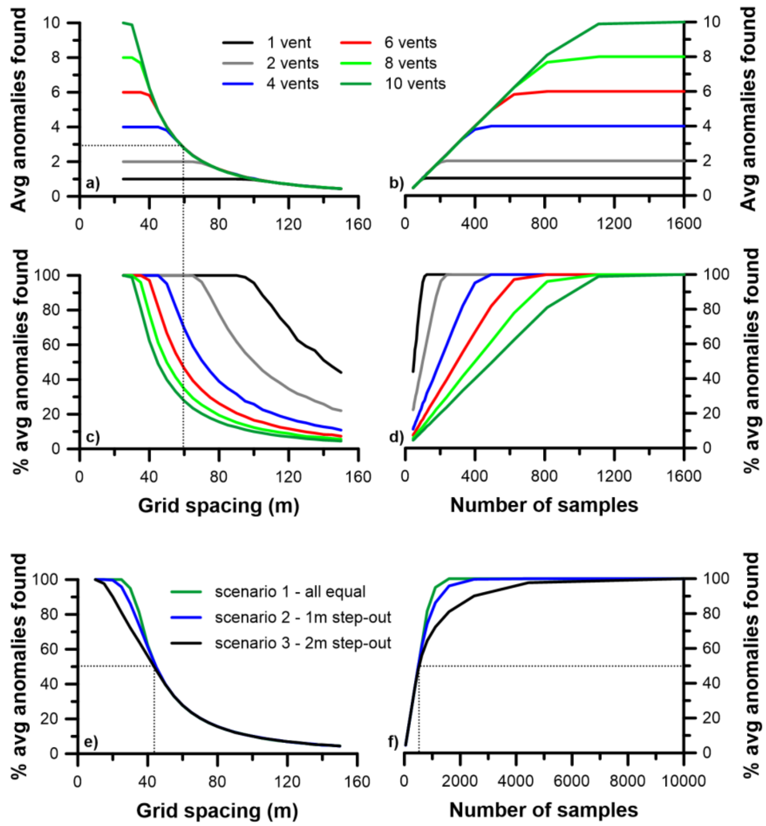

Figure 4. The absolute number (a, b) and percent (c, d) of anomalies found with square grid sampling over a range of grid spacings (a, c) and number of samples (b, d) for 6 scenarios having a different number of equal-sized anomalies whose sum is always 10,000m^2. The same for three scenarios having 10 anomalies of different sizes (e, f); note the expanded X scale in (f).

3.2. Quantifying leakage

3.2.1. Background flux estimates

As stated above, the background biogenic CO_2 flux must be estimated and subtracted from the total measured flux to calculate the leakage flux rate. Various approaches have been used in the literature.

For example, the average background value can be calculated using measurements from a similar, sub-set area where no leakage occurs. To assess this the 1km^2 synthetic background dataset was subdivided into 16, 250 x 250 m sub-areas (Figure S3b). MC-Flux was used to create 1000 sub-sampling files for each sub-area (random grid with 35 m spacing) as well as the total area (random grid with 140m spacing). Each realization yielded about 50 samples which were used to calculate the average and MVUE flux rates. Although the statistical distribution of the estimates from each sub-area are relatively narrow, there is a wide range of values (blue boxes in

Figure 5a) compared to that for the entire area (red box in Figure 5a) due to the non-random distribution of the synthetic data. Although some sub-areas yield median estimates that are similar to the true value of 20 g m^{-2} d^{-1}, others vary by as much as ± 4 g m^{-2} d^{-1}.

Another approach uses a log-normal probability plot of all survey data to separate and characterize the background. To examine this approach the HR Latera dataset (Figure S6) was subsampled 500 times with the offset grid approach using three different spacings (20, 30, 40m). Log-normal probability plots were made for 50 realizations of each of the three datasets, upper background thresholds were estimated graphically based on the inflection point in the data, and the MVUE of the background population was calculated for each realization. Different inflection points (e.g., Figure 5b) and sampled background values lead to background flux estimates (Figure 5c) that vary by up to a maximum of ± 3 g m^{-2} d^{-1} for the wider spacings.

A third approach, similar to the previous but with background flux calculated using only the area that it is estimated to occupy, has not been assessed because it is impractical for MC simulations. Based on the results above, however, it is likely that this approach would have a similar level of uncertainty. A fourth method, using isotopes or co-migrating gases (e.g., Bini et al., 2020), should instead have significantly lower uncertainties.

Although the true background flux uncertainty will be site-specific, and will depend both on the approach used and researcher experience, these results illustrate the potential for errors on the order of a few g m^{-2} d^{-1} for the estimated average. While small, it could have an important impact on the final leakage flux value if the survey area is large and/or the average biogenic flux is proportionally high. As an example, given a 100,000 m^2 area with a total flux of 4 t d^{-1} and a true average background flux of 20 g m^{-2} d^{-1}, an estimate that is ± 2 g m^{-2} d^{-1} from this value will result in an error of about ±10% in the final leakage estimate. Finally it should be remembered that these examples only consider uncertainty related to statistical sampling, and do not take into account uncertainty caused by temporal variability (given that flux surveys are conducted over many hours during the day and often over multiple days).

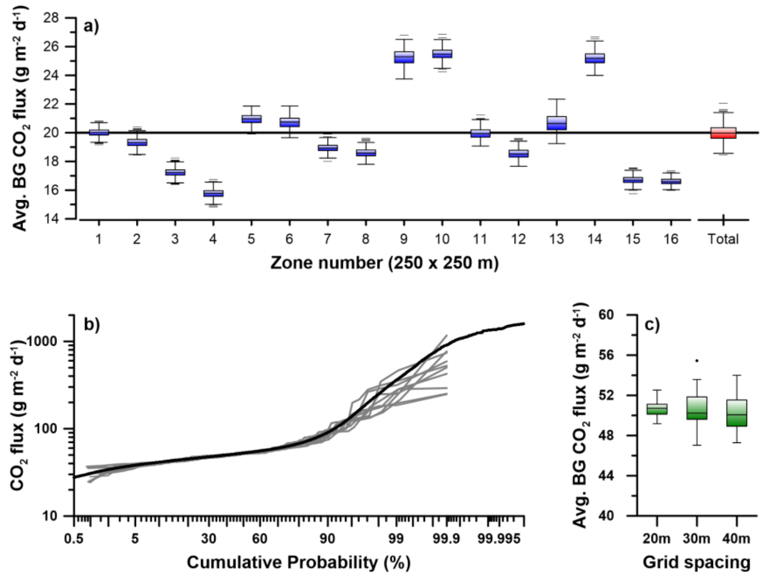

Figure 5. (a) Range of estimated average background CO_2 flux for the synthetic dataset, including the total 1km^2 area (red) and 16 sub-set areas (blue), compared to the true average (horizontal line). (b) Log-normal probability plots of the high density Latera dataset (black line) and the first 10 subsampling realizations at 30m sample spacing (grey lines). (c) Range of the estimated average background flux for the Latera data, based on log-normal probability plot interpretation of 50 subsample datasets at 20, 30, and 40m offset grid spacing.

3.2.2. Leakage flux estimates - calculation approaches

All leakage flux simulations discussed in this and the following sections were performed using the HR Latera data as input (Figure S6). For this section a total of 500 sub-sampling realizations were performed with MC-Flux using the off-set grid and random strategies at 8 different sample densities each (5, 10, 15, 20, 25, 30, 35, 40m grid spacing, or equivalent number of samples); these strategies were chosen because they yielded the highest and lowest probability of finding anomalies, respectively, in Section 3.1. Total flux of the resultant sub-sampled datasets was calculated using all four methods. Certain parameters were fixed using the "true" values of the complete HR dataset to standardize the processing of thousands of datasets. For both kriging and SGS the standardized variogram model was defined as 0.02 Nugget + 0.98 * Spherical(37) (Figure S6d) and the search radius set to 46m, which is slightly larger than the variogram range to guarantee sufficient points for the widest grid spacing. Parameters for back-transforming the normal scores in sGsim were those used to create the original HR dataset (i.e., lower tail linear extrapolation to 20 and upper tail hyperbolic extrapolation to 4000 with $\omega = 2$). Other parameters fixed for the sGsim calculations include: previously simulated nodes to use =

16; multiple refinement grids = 0; minimum and maximum data for simulation = 0 and 8; kriging type = simple. Finally, the average background flux calculated in Section 2.1 (51 g m^{-2} d^{-1}) was subtracted across the entire grid to convert the calculated total flux to leakage flux. These simplifications remove any uncertainties related to these parameters, thus highlighting variations that are primarily related to sub-sampling effects. A comparison of estimates made for 5 datasets from 3 grid sizes using these standardized parameters versus those calculated individually and manually for each file (Figures S17, S18) show relatively small changes and indicate that the approach is valid (Figure S19). Box plots for all data are given in Figure S20, while a subset of the offset grid simulations is given in Figure 6.

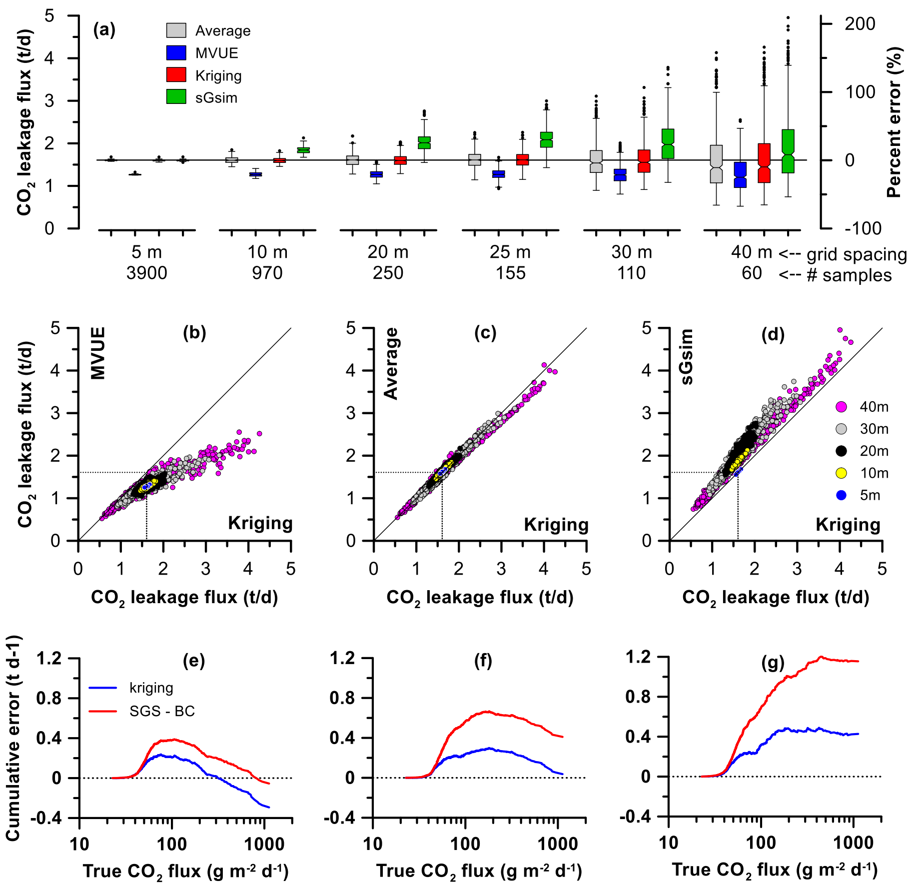

Figure 6. CO_2 leakage flux estimates of the Latera HR data using offset grid sub-sampling and four different calculating methods. (a) Statistical distribution versus the true value (horizontal line). (b-d) Comparison of kriging estimates with those defined using the average (b), MVUE (c), and SGS (d) calculation methods; the dashed lines indicate the true value. (e-f) Cumulative error plots for the 20m grid sub-sampled files that yielded the lowest (e), median (f) and highest (g) SGS estimates.

The MVUE approach consistently under-estimates the leakage flux rates (Figure 6a), even at the 5 m grid spacing, likely because this approach is only valid for a single log-normal population (Schroder et al. 2017). The greater scatter of the MVUE data relative to kriging (Figure 6b) is likely due to the variable effect that different leakage samples have on the final calculated value. Although some research has indicated that the MVUE approach may be valid for CO_2 leakage estimates if the data is log-normally distributed (Lewicki et al., 2005; Elio et al., 2016), these results appear to support the affirmation by Schroder et al. (2017) that it is not appropriate. There is a sharp contrast between the MVUE and AM trends, as also shown by a change in their ratio when using only background values versus including progressively more leakage values (Figure S21). The AM data distribution is often centered on the true leakage flux value (Figure 6a) and has a distribution similar to that kriging (Figure 6c), despite the fact that it is also only valid for a single, in this case normal, population.

The generally linear relationship between kriging and SGS (Figure 6d) is expected, as the latter has kriging at its core and the more realizations performed the more smoothed and similar to the former the results become (Cardellini et al. 2003). However, for some grid spacings (10 to 30 m) SGS on average overestimates CO_2 leakage (Figure 6a) and the slope between kriging and SGS deviates slightly from 1:1 (Figure 6d). Cumulative error plots (i.e., estimated minus true values, summed for increasing true values) for the 20m offset grid subsampled files that gave the lowest (Figure 6e), median (Figure 6f), and highest (Figure 6g) total flux estimates using SGS show how both methods generally overestimate the low values and underestimate the high values, which is a well-known artefact of kriging (Cardellini et al. 2003). However, while their trends are very similar at the low end, SGS continues to over-estimate across a mid-range of values before once again paralleling the kriging trend. These plots show that the final estimate for both methods is a balance between over- and under-estimating across the range of values, rather than a true representation, and how the higher SGS estimates are caused primarily by over-estimation in mid-range values.

To better understand the SGS results, sensitivity analyses were performed with MC-Flux using 100 sub-sampled datasets (offset grid, 20 m spacing) as input, systematically changing various SGS parameters. Base case (BC) calculations were made for AM, kriging and SGS (Figure 7a), with SGS BC parameters set to those used to produce the data in Figure 6.

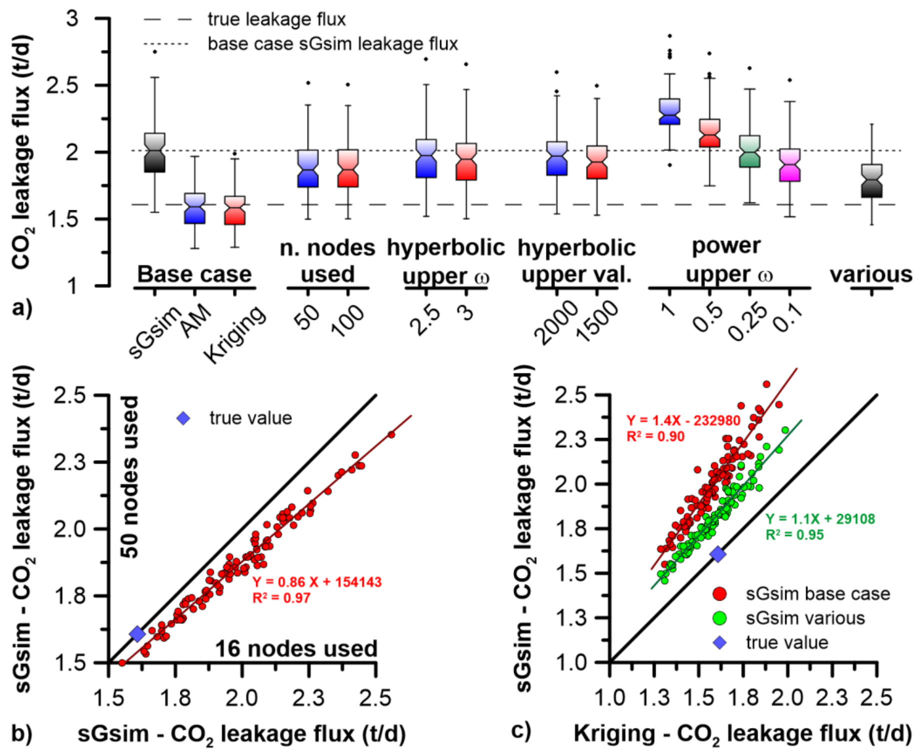

Figure 7. (a) Statistical distribution of sensitivity analyses of various sGsim parameters; see text for description. Comparison of sGsim-estimated leakage flux using 16 nodes (base case) and 50 nodes (b) and base case and “various” sGsim-estimated leakage flux versus the corresponding kriging results (c).

The sGsim algorithm uses a large, odd integer as a seed to the pseudorandom number generator, and each realization can be reproduced exactly by re-running the simulation algorithm using that seed. In all simulations presented thus far the default value of 69069 has been used. For this test, the two sub-sample files that yielded the highest and lowest leakage flux estimates were each processed 100 times with sGsim, each time performing the usual 100 realizations with the BC parameters but with a unique random seed value. For both input files the statistical distribution of the leakage estimates were defined by $1\sigma \approx 1.5\%$ and a total (i.e., minimum to maximum) range of about 8%. This uncertainty, however, is random and does not cause a fixed upwards or downwards shift in the data, as seen with other parameters.

As each new node is simulated, sGsim adds this value to the original dataset used for kriging of subsequent points, leading to progressively more values and slower computations. To speed up the algorithm the user can limit the number of previously simulated nodes that are used at each step. Sensitivity analyses show a drop in the median SGS leakage flux estimate by about

0.14 t d^{-1} (ca. 7%) when the number of nodes is increased from 16 (BC) to 50, but with little subsequent change using 100 nodes (Figure 7a). Both 50 and 100 nodes produce slightly narrower distributions, with greater impacts observed for those sub-sampled files that yield higher leakage flux estimates (Figure 7b). These results are most likely due to the observation that the use of a moving neighborhood of conditioning values can result in sGsim realizations that poorly reproduce second-order statistics, leading to simulated variograms that are biased relative to the theoretical one (Paravarzar et al. 2015). For example, Emery (2004) showed that using 20 nodes resulted in an increase in the apparent variogram range by up to 25%, while results with 100 nodes were better but still produced an increase of 14%. In the context of this work, an overestimated range would extend leakage anomalies over larger areas (compared to kriging) and yield a larger CO_2 leakage estimate.

If input data are not normally distributed they must be log-transformed prior to sGsim calculations and then back-transformed for final output, which requires the selection of the extrapolation model type and extreme values for both the lower and upper dataset tails. While lower tail characteristics can be estimated relatively accurately, the upper tail is much more difficult to assess. Estimating the maximum value by extrapolating the CPP is a valid approach (e.g., Cardellini et al. 2003), however the range of potential estimates that could result from probabilistically equivalent sub-sample datasets, like those shown in Figure 5b, indicates the uncertainty in this approach. Initial tests involved changing parameters of the hyperbolic model (Figure 7a). It was found that increasing the ω parameter (related to trend curvature) from 2 (BC) to 2.5 to 3 decreased the median estimate by 1.8% and 3.2%, respectively, while decreasing the maximum value from 4000 g m^{-2} d^{-1} (BC) to 2000 and then 1500 (actual maximum value in the dataset) decreased it by 2.0 and 4.3%, respectively. Modification of the ω parameter for the power model had even larger effects, with a value of 1 (equivalent to the linear model) giving a median value that is 13% above the BC while a value of 0.1 yielded one that was about 5.2% below (Figure 7a).

Based on the above results a final test ("various" in Figure 7) was performed that changed three parameters at the same time (50 nodes used and hyperbolic model with a maximum value of 1500 and $\omega = 2.5$). This resulted in a narrowing of the distribution and decrease of the median value by 0.22 t d^{-1} (c. 11%). Plotted against the kriging results this final test yielded a higher R^2 and a slope closer to 1 compared to the base case (Figure 7c). Despite these improvements, the sGsim still overestimated, on average, the true value by about 0.2 t d^{-1} for these tests on the 20 m offset grid sub-samples. Two possible explanations can be inferred from Emery (2004). First, a continued overestimate of the variogram range is possible despite the increased number of retained nodes. Second, the simulated area may be too small, in that the ratio between variogram range and domain length is 0.14 and 0.1 in the x and y directions, respectively, compared to the recommended value of <0.05.

3.2.3. Leakage flux estimates - accuracy

Up to 3000 MC-Flux sub-sample realizations were performed using the same input, assumptions and sample densities described above for the results in Figure 6, but applying all five sampling strategies (note that the number of unique square, offset, and triangular grid realizations are limited by (grid spacing)^2). Considering the similar behavior of AM and kriging described in the previous section, kriging and sGsim simulations were not performed.

The resultant AM data were used to calculate the probability that a given number of samples would yield a certain level of accuracy relative to the true value (Figure 8). Results of the random sampling strategy (Figure 8a) show smooth trends for all considered accuracy levels and low probability of high accuracy even for large number of samples, in agreement with that observed by Wong (2018). For example, there is only a 56% chance that 1000 samples would yield a leakage flux estimate that is within 10% of the true value while the same number of samples have an 88% chance of being within 20%. Unlike Wong (2018), however, we have also examined the response of different gridded sampling strategies, which show markedly improved results. For example, the offset grid data (Figure 8b) show trends that rise much more rapidly with far fewer samples. These trends are more irregular than those for the random sampling, likely due to the effect of sample node spacing combined with the individual size of each anomaly and the average distance between them.

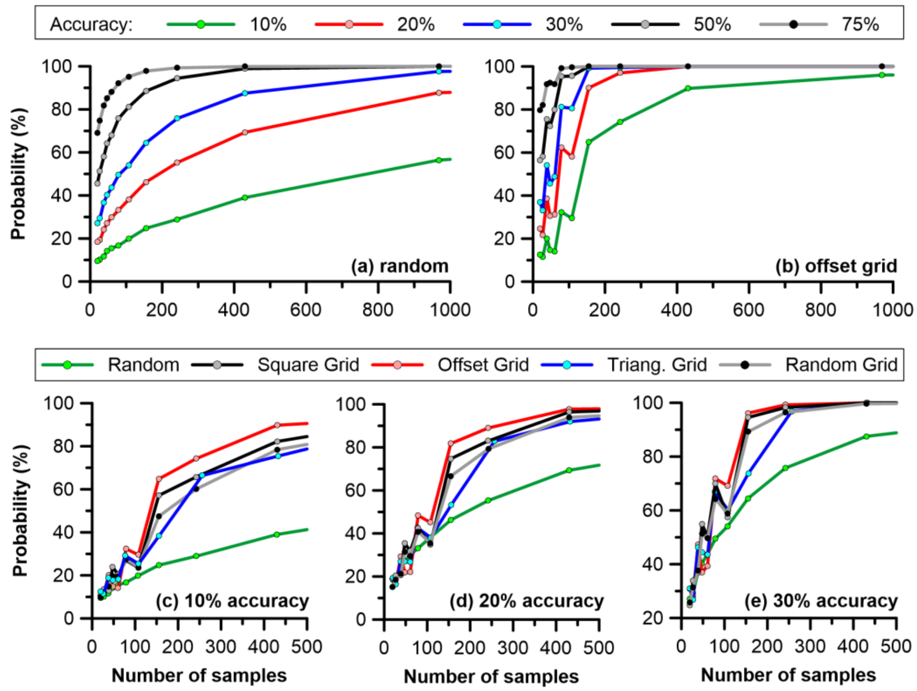

Figure 8. Plots showing the probability that a certain number of samples will produce a given level of accuracy (lines) for random (a) and offset grid (b) sub sampling, and for all five sub-sampling strategies at 10% (c), 20% (d), and 30% (e) levels of accuracy.

A direct comparison of the results for the different sampling strategies at accuracy levels of 10% (Figure 8c), 20% (Figure 8d), and 30% (Figure 8e) show that random sampling consistently provides the poorest results, the different gridded approaches show similar trends (with triangular grid variations likely due to larger sample spacing for an equivalent number of samples, see Figure S7), and the offset grid yields the most accurate results (in agreement with results presented in Section 3.1). From 20 to 100 samples (i.e., 70 to 30 m grid spacing, equivalent to 200 to 1000 samples km-2) the trends of all 5 strategies are relatively similar, with a less than 30% chance of 10% accuracy, less than 40% chance of 20% accuracy, and less than 60% chance of 30% accuracy. Instead at the next sample density (i.e, 155 samples, 25 m grid spacing, 1550 samples km^{-2}) there is a rapid improvement in the grid sampling results, with probabilities doubling in some cases. To put these results in context, the use of the offset grid strategy and a desired 90% probability level would require 15m spacing to obtain 10% accuracy, 20m for 15%, 25m for 20% and 30m for 40%. It should be acknowledged that these results are based on an original dataset having a ca. 10 m sample spacing, and although sGsim processing was performed to reduce bias, a link cannot be excluded. For this reason similar simulations using high density datasets from other sites should be conducted in the future.

4. Conclusions

Although it is true that each site is unique and its associated flux data must be interpreted individually to obtain the best results (Elío et al. 2016), the Monte Carlo approach used here (with its necessary standardization of input variables) provides useful information that can help reduce uncertainties and errors in the soil flux method.

The offset and triangular grid sample strategies are recommended due to their superior performance under all conditions except low sample densities, where they were equivalent to the other methods. However, because the orientation of elongated anomalies were found to have a large impact on probabilities, due to alignment within grid gaps, any available site information regarding anisotropy, such as air photos, should be taken into consideration for deciding the orientation of the grid itself.

Although sequential Gaussian simulation is the most commonly used method for interpolating and quantifying leakage flux data, the results presented here show how its estimate is sensitive to the chosen input parameters and, at least for the dataset used, slightly overestimated for mid-range sample spacings. In contrast, the much simpler and less subjective approaches using the arithmetic mean or kriging yielded probabilistic distributions were centered on the true value and thus may be more appropriate (although kriging accuracy seems to be linked to how well over- and under-estimated values are balanced across the grid). Additional simulations using other high resolution flux datasets from other sites should be performed to confirm these results.

The need for large numbers of closely spaced samples to accurately define leakage flux is well known, and the results presented here help to quantify the potential level of uncertainty that can be expected at various sample densities. This raises the question of whether there are other approaches that could yield the same or lower uncertainty levels but with fewer samples, thus freeing resources for other analyses to better separate background and leakage populations (e.g., Bini et al. 2020). In this regard, both kriging and SGS would benefit from the use of a high resolution dataset of a linked parameter that could be used to both help choose appropriate sample spacing/ locations and for variogram definition and co-kriging. One such possibility is a map of CO_2 concentration anomalies at ground level (Beaubien et al. 2018) or just above it

(Barkwith et al. 2020), data which can be collected rapidly at high resolution. Our group is in the process of assessing this approach by combining these two datasets for MC-Flux simulations.

Acknowledgments

The assistance of Pietro Sacco and Davide de Angelis for the collection of the Latera field data is gratefully acknowledged.

Declarations

Funding: The research leading to these results received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653718.

Conflicts of interest/Competing interests: The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material: Datasets and configuration files used to conduct the reported simulations are available from the Zenodo open-source online repository at https://doi.org/10.5281/zenodo.4573725 (Beaubien and Bigi, 2021b).

Code availability: The MC-Flux (V1.0) installation package and user's manual are available from the Zenodo open-source online repository at https://doi.org/10.5281/zenodo.4573575 (Beaubien and Bigi, 2021a).

Authors' contributions: Conceptualization [SEB, SL], Methodology, Software, Visualization and Writing – Original Draft [SEB]; Formal Analysis [SEB, GC, MGF]; Validation [GC, MGF]; Writing – Review and Editing [GC, MGF, SB]; Supervision [SL]; Project administration and Funding acquisition [SB, SL].

References

Annunziatellis A, Beaubien SE, Bigi S, Ciotoli G, Coltella M, Lombardi S (2008) Gas migration along fault systems and through the vadose zone in the Latera caldera (central Italy): Implications for CO2 geological storage. Int J Greenhouse Gas Control, 2/3: 353-372. https://doi.org/10.1016/j.ijggc.2008.02.003

Bain WG, Hutyra L, Patterson DC, Bright AV, Daube BC, Munger JW, Wofsy SC (2005) Wind-induced error in the measurement of soil respiration using closed dynamic chambers. Agricultural and Forest Meteorology, 131(3): 225-232. https://doi.org/10.1016/j.agrformet.2005.06.004

Barkwith A, Beaubien SE, Barlow T, Kirk K, Lister TR, Tartarello MC, Taylor-Curran H (2020) Using near-surface atmospheric measurements as a proxy for quantifying field-scale soil gas flux. Geosci Instrum Method Data Syst, 9(2): 483-490. https://doi.org/10.5194/gi-9-483-2020

Beaubien SE (2015) The mapping and quantification of CO2 leakage and its potential impact on groundwater quality. Ph.D. Thesis, Università Ca' Foscari, Venezia, Italia, 163 pp. http://hdl.handle.net/10579/6509

Beaubien SE, Annunziatellis A, Ciotoli G, Lombardi S (2009) Near Surface Gas Simulator (NSGS): A Visual Basic program to improve the design of near-surface gas geochemistry surveys above CO2 geological storage sites, 6th European Geosciences Union General Assembly 2009, Vienna, Austria. https://meetingorganizer.copernicus.org/EGU2009/EGU2009-12489.pdf

Beaubien SE, Bigi S (2021) Example data and configuration / input files for conducting probability and accuracy simulations of CO2 flux surveys using the MC-Flux program. https://doi.org/10.5281/zenodo.4573725

Beaubien SE, Bigi S (2021) MC-Flux (Version 1.0): a Monte Carlo-based program to determine how sample spacing, sample strategy, and calculation approaches impact the probability of finding CO2 leakage anomalies and the accuracy of total CO2 flux estimates. https://doi.org/10.5281/zenodo.4573575

Beaubien SE, Ciotoli G, Coombs P, Dictor MC, Krüger M, Lombardi S, Pearce JM, West JM (2008) The impact of a naturally-occurring CO2 gas vent on the shallow ecosystem and soil chemistry of a Mediterranean pasture (Latera, Italy). Int J Greenhouse Gas Control, 2/3: 373-387. https://doi.org/10.1016/j.ijggc.2008.03.005

Beaubien SE, Jones DG, Goldberg T, Barkwith KAP, Bigi S, Graziani S, Kirk K, Mattei E, Mulder B, Pettinelli E, Ruggiero L, Tartarello M-C (2018) Innovative tools for rapidly mapping / quantifying CO2 leakage and determining its origin, 14th Greenhouse Gas Control Technologies Conference (GHGT-14) Melbourne, Australia. https://dx.doi.org/10.2139/ssrn.3366264

Bini G, Chiodini G, Lucchetti C, Moschini P, Caliro S, Mollo S, Selva J, Tuccimei P, Galli G, Bachmann O (2020) Deep versus shallow sources of CO2 and Rn from a multi-parametric approach: the case of the Nisyros caldera (Aegean Arc, Greece). Scientific Reports, 10(1): 13782. https://doi.org/10.1038/s41598-020-70114-x

Bond-Lamberty B, Pennington SC, Jian J, Megonigal JP, Sengupta A, Ward N (2019) Soil respiration variability and correlation across a wide range of temporal scales. Journal of Geophysical Research: Biogeosciences, 124(11): 3672-3683. https://doi.org/10.1029/2019jg005265

Caers J (2002) Direct sequential indicator simulation. In: W. Kleingeld and D. Krige (Editors), 6th International Geostatistics Congress, South Africa, Cape Town, South Africa, pp. 39-48. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.217.7285andrep=rep1andtype=pdf

Carapezza ML, Granieri D (2004) CO2 soil flux at Vulcano (Italy): comparison between active and passive methods. Applied Geochemistry, 19: 73-88. https://doi.org/10.1016/S0883-2927(03)00111-2

Cardellini C, Chiodini G, Frondini F (2003) Application of stochastic simulation to CO2 flux from soil: Mapping and quantification of gas release. Journal of Geophysical Research: Solid Earth, 108(B9): 2425. http://dx.doi.org/10.1029/2002JB002165

Chiodini G, Baldini A, Barberi F, Carapezza ML, Cardellini C, Frondini F, Granieri D, Ranaldi M (2007) Carbon dioxide degassing at Latera caldera (Italy): Evidence of geothermal reservoir and evaluation of its potential energy. Journal of Geophysical Research, 112(B12204). https://doi.org/10.1029/2006JB004896

Chiodini G, Cardellini C, Caliro S, Avino R, Donnini M, Granieri D, Morgantini N, Sorrenti D, Frondini F (2020) The hydrothermal system of Bagni San Filippo (Italy): fluids circulation and CO2 degassing. Italian Journal of Geosciences, 139(3): 383-397. https://doi.org/10.3301/IJG.2020.12

Chiodini G, Cioni R, Guidi M, Marini L, Raco B (1998) Soil CO2 flux measurements in volcanic and geothermal areas. Applied Geochemistry, 13: 543-552. https://doi.org/10.1016/S0883-2927(97)00076-0

Davidson JR, 1995. Monte Carlo tests of the Elipgrid-PC algorithm. ORNL/TM-12899. Oak Ridge National Laboratory, Oak Ridge, Tennessee. https://doi.org/10.2172/52637

Deutsch CV, Journel AG (1997) GSLIB: Geostatistical software library and users guide. Oxford Univ. Press, New York, USA, 2nd Edition, 369 pp. https://doi.org/10.1080/00401706.1995.10485913

Elío J, Ortega MF, Chacón E, Mazadiego LF, Grandia F (2012) Sampling strategies using the "accumulation chamber" for monitoring geological storage of CO2. Int J Greenhouse Gas Control, 9: 303-311. https://doi.org/10.1016/j.ijggc.2012.04.006

Elío J, Ortega MF, Nisi B, Mazadiego LF, Vaselli O, Caballero J, Chacón E (2016) A multi-statistical approach for estimating the total output of CO2 from diffuse soil degassing by the accumulation chamber method. Int J Greenhouse Gas Control, 47: 351-363. https://doi.org/10.1016/j.ijggc.2016.02.012

Emery X (2004) Testing the correctness of the sequential algorithm for simulating Gaussian random fields. Stochastic Environmental Research and Risk Assessment, 18(6): 401-413. https://doi.org/10.1007/s00477-004-0211-7

Gilardi N, Bengio S, Kanevski M (2002) Conditional Gaussian mixture models for environmental risk mapping, Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing, pp. 777-786. https://doi.org/10.1109/NNSP.2002.1030100

Gilbert RO (1987) Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold Company Inc., New York, 332 pp. ISBN: 978-0-471-28878-7

Hood GM, 2010. PopTools version 3.2.5. http://www.bioquest.org/esteem/esteem_details.php?product_id=248

Kutzbach L, Schneider J, Sachs T, Giebels M, Nykanen H, Shurpali NJ, Martikainen PJ, Alm J, Wilmking M (2007) CO2 flux determination by closed-chamber methods can be seriously biased by inappropriate application of linear regression. Biogeosciences, 4(6): 1005-1025. https://bg.copernicus.org/articles/4/1005/2007/bg-4-1005-2007.pdf

Leon E, Vargas R, Bullock S, Lopez E, Panosso AR, La Scala N (2014) Hot spots, hot moments, and spatio-temporal controls on soil CO2 efflux in a water-limited ecosystem. Soil Biology and Biochemistry, 77: 12-21. https://doi.org/10.1016/j.soilbio.2014.05.029

Lewicki JL, Bergfeld D, Cardellini C, Chiodini G, Granieri D, Varley N, Werner C (2005) Comparative soil CO2 flux measurements and geostatistical estimation methods on Masaya volcano, Nicaragua. Bull Volcanol, 68: 76-90. https://doi.org/10.1007/s00445-005-0423-9

Matzke BD, Wilson JE, Newburn LL, Dowson ST, Hathaway JE, Sego LH, Bramer LM, Pulsipher BA (2014) Visual Sample Plan: Version 7.0 User’s Guide. Pacific Northwest National Laboratory for the U.S. Department of Energy. Richland, Washington, USA., pp. 291. https://doi.org/10.2172/1124046

Oldenburg CM, Lewicki JL, Hepple RP, 2003. Near-surface monitoring strategies for carbon dioxide storage verification, Lawrence Berkeley National Laboratory, Report LBNL-54089. https://escholarship.org/uc/item/1cg241jb

Paravarzar S, Emery X, Madani N (2015) Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. Comptes Rendus Geoscience, 347(2): 84-93. https://doi.org/10.1016/j.crte.2015.05.008

Pettinelli E, Beaubien SE, Zaja A, Menghini A, Praticelli N, Mattei E, Di Matteo A, Annunziatellis A, Ciotoli G, Lombardi S (2010) Characterization of a CO2 gas vent using various geophysical and geochemical methods. Geophysics, 75(3): B137-B146. http://dx.doi.org/10.1190/1.3420735

Riveros-Iregui DA, McGlynn BL (2009) Landscape structure control on soil CO2 efflux variability in complex terrain: Scaling from point observations to watershed scale fluxes. Journal of Geophysical Research: Biogeosciences, 114(G2). https://doi.org/10.1029/2008jg000885

Sainju UM, Jabro JD, Stevens WB (2008) Soil carbon dioxide emission and carbon content as affected by irrigation, tillage, cropping system, and nitrogen fertilization. Journal of Environmental Quality, 37(1): 98-106. https://doi.org/10.2134/jeq2006.0392

Schroder IF, Wilson P, Feitz AF, Ennis-King J (2017) Evaluating the performance of soil flux surveys and inversion methods for quantification of CO2 leakage. Energy Procedia, 114: 3679-3694. https://doi.org/10.1016/j.egypro.2017.03.1499

Singer DA (1972) ELIPGRID, a FORTRAN IV program for calculating the probability of success in locating elliptical targets with square, rectangular, and hexagonal grids. Geocom Programs, 4: 1-16. https://www.researchgate.net/publication/236410051

Singer DA (1975) Relative efficiencies of square and triangular grids in the search for elliptically shaped resource targets. J Research U.S. Geological Survey, 3(2): 163-167. https://pubs.er.usgs.gov/publication/70164418

Viveiros F, Chiodini G, Cardellini C, Caliro S, Zanon V, Silva C, Rizzo AL, Hipólito A, Moreno L (2020) Deep CO2 emitted at Furnas do Enxofre geothermal area (Terceira Island, Azores archipelago). An approach for determining CO2 sources and total emissions using carbon isotopic data. Journal of Volcanology and Geothermal Research, 401: 106968. https://doi.org/10.1016/j.jvolgeores.2020.106968

Wong CLY (2018) Analysis of the number of flux chamber samples and study area size on the accuracy of emission rate measurements. Journal of the Air and Waste Management Association, 68(10): 1103-1117. https://doi.org/10.1080/10962247.2018.1469555