



Environmental Analysis Program

Pomona College
Biogeochemistry Laboratory

SOP-F005

Random Sampling Design

Standard Operating Procedure

SOP Number	SOP-F005
Revision	1.1
Author(s)	Marc Los Huertos
Date	2026-06-26
Organization	EA Program, Pomona College

FOR INTERNAL USE ONLY

This document is the property of the Environmental Analysis Program at Pomona College.
Unauthorized distribution is not permitted.

Table of contents

1 Purpose	1
2 Scope	1
3 Responsibilities	1
4 Safety and Hazards	1
4.1 Required PPE	1
4.2 Emergency Procedures	2
5 Definitions	2
6 Required Materials	2
7 Biases and Interferences	3
8 Procedure	3
8.1 Generating Random Numbers	3
8.2 Simple Random Sampling	4
8.3 Grid-Based Random Sampling (Field)	4
8.3.1 Establishing the Baseline Grid	4
8.3.2 Generating Coordinates and Placing Quadrats	5
8.3.3 Recording Data Within Quadrats	5
8.4 Stratified Random Sampling	6
8.5 Map-Based Sampling with Exclusion Zones	6
9 Data Management	6
10 References	7
11 Related SOPs	7
12 Training Record	7

1. Purpose

This SOP provides standardized procedures for designing and executing random sampling in field and laboratory settings. Random sampling ensures that every unit in the target population has an equal and independent probability of selection, eliminating researcher bias and supporting valid statistical inference (Elzinga et al., 1998; Krebs, 1999).

2. Scope

- **Applies to:** Any EA Program study where population estimates are the goal, including vegetation surveys, soil sampling, water quality monitoring, and laboratory subsampling
- **Does not apply to:** Targeted or judgmental sampling designs (e.g., hot-spot investigations); mark-release-recapture studies for mobile organisms (see separate SOP)
- **Sampling types covered:** Simple random sampling, grid-based random sampling, and stratified random sampling

3. Responsibilities

Role	Responsibility
PI / Field Lead	Approve sampling design, review sampling plan
Trained Personnel	Execute random selection, collect and record data
New Students	Complete training before independent fieldwork

4. Safety and Hazards

Hazards

Field sampling involves site-specific hazards including uneven terrain, heat exposure, wildlife encounters, and sharp equipment (stakes, pins). Review site conditions before each sampling event.

4.1. Required PPE

- ☐ Closed-toe shoes or boots
- ☐ Sun protection (hat, sunscreen)
- ☐ Gloves (when handling equipment or samples)

4.2. Emergency Procedures

Refer to [SOP-F001 Field Safety](#). Carry a first aid kit and charged cell phone to all field sites.

5. Definitions

- **Border rule:** A convention for counting organisms that overlap a quadrat edge; the standard rule is to count organisms touching the top and right edges and exclude those touching the bottom and left edges; must be applied consistently within a study to avoid double-counting (Elzinga et al., 1998).
- **Grid-based random sampling:** A method in which a coordinate grid is laid over the study area and quadrat locations are selected using randomly generated X and Y coordinates; the standard approach for quadrat-based ecological surveys (Elzinga et al., 1998).
- **Power analysis:** A calculation to determine the minimum sample size (n) needed to detect a biologically meaningful effect with acceptable Type I error (α) and Type II error (β) probability; should be completed before fieldwork begins (Krebs, 1999).
- **Quadrat:** A square frame of known area (typically 0.25, 0.5, or 1.0 m²) used to delineate sampling units in the field (Elzinga et al., 1998).
- **Sampling frame:** A numbered list or spatial grid that enumerates every unit in the population; units missing from the frame have zero probability of selection and introduce bias (Krebs, 1999).
- **Sampling universe (population):** The complete set of units from which a sample will be drawn; must be explicitly defined before generating a sampling frame (Krebs, 1999).
- **Simple random sampling:** A design in which n units are selected at random from a numbered list (sampling frame) such that every possible combination of n units has an equal probability of selection; appropriate when the population is homogeneous (Krebs, 1999).
- **Stratified random sampling:** A design in which the population is divided into internally homogeneous subgroups (strata) and random samples are drawn independently from each stratum; reduces variance compared to simple random sampling when strata differ from each other (Elzinga et al., 1998; Krebs, 1999).
- **Stratum:** A subdivision of the sampling universe into zones that are internally homogeneous (e.g., by elevation band, vegetation type, or soil moisture class); the basis for stratified random sampling.

6. Required Materials

Item	Notes
Two survey tape measures (20–50 m)	For establishing coordinate grids
Ground pins or stakes	To anchor tape measures
Quadrat frame(s)	Size appropriate to target organisms
Random number generator app or R	See §8.1
GPS unit	To record sampling locations
Field data sheets and waterproof markers	Pre-printed with site and date fields
Compass	For establishing grid orientation
Field identification guides	If sampling vegetation or organisms

7. Biases and Interferences

Representative sampling requires strict adherence to randomized procedures. Common sources of bias include:

- **Incomplete sampling frames.** Any unit missing from the frame has zero probability of selection (Krebs, 1999).
- **Haphazard placement.** Never “throw” a quadrat to choose a location; human throws cluster toward the center of the area, violating equal probability assumptions (Elzinga et al., 1998).
- **Environmental heterogeneity.** If the study area contains distinct gradients, simple random sampling may under-represent one end. Use stratified random sampling instead (§8.4).
- **Edge effects.** Establish and apply a consistent border rule when counting organisms in quadrats (§8.3).

8. Procedure

8.1. Generating Random Numbers

Use a software-based random number generator to produce unbiased selections. Never rely on manual or mental number selection (Krebs, 1999).

In R, generate a set of random integers within a defined range:

```
# Generate 15 random integers between 1 and 100
set.seed(NULL) # ensure a fresh seed each session
random_vals <- sample(1:100, size = 15, replace = FALSE)
sort(random_vals)
```

```
[1] 3 9 12 15 20 35 44 51 59 63 66 75 78 84 100
```

For field use without a computer, a printed random number table is acceptable. Begin at a pre-selected (arbitrary) starting position and read values sequentially. Table 1 provides a reference table of 200 random numbers between 0 and 1000.

Table 1: Reference table of 200 random numbers (0–1000)

1	2	3	4	5	6	7	8	9	10
924.2	634.1	556.0	247.2	144.7	912.8	778.9	886.8	4.8	373.9
537.2	379.9	597.8	374.3	856.6	972.9	414.4	848.9	367.1	208.8
469.2	942.1	197.0	595.3	509.7	56.6	574.1	244.4	943.6	610.2
288.6	755.0	79.2	751.2	920.4	192.6	399.1	194.8	98.3	188.4
770.1	227.6	971.5	722.2	232.8	578.3	91.7	278.5	457.4	21.5
772.8	914.7	509.7	470.0	195.0	28.6	174.7	224.9	523.8	710.5
39.3	620.4	49.7	176.0	152.5	491.3	525.4	49.3	529.7	164.2

Table 1: Reference table of 200 random numbers (0–1000)

1	2	3	4	5	6	7	8	9	10
434.9	319.1	675.4	111.4	165.3	589.5	345.6	973.2	740.5	826.2
472.2	76.3	776.3	616.9	367.0	55.5	827.1	965.7	508.6	166.7
273.8	260.8	753.4	667.2	85.9	0.5	638.3	696.6	490.0	417.1
674.3	883.8	524.8	999.1	866.6	152.6	594.7	998.7	515.1	667.4
2.0	17.5	883.5	690.5	647.1	48.2	198.4	445.5	535.9	512.7
25.1	314.9	348.6	152.1	519.7	942.0	314.5	224.7	833.1	369.5
432.1	211.8	638.2	878.3	543.3	179.5	228.5	574.9	528.1	357.5
499.4	906.2	737.8	973.9	837.4	931.0	700.3	439.8	755.1	90.1
388.7	650.0	287.8	373.8	126.8	600.6	194.7	755.6	212.7	156.0
395.4	129.3	192.8	390.5	820.9	411.9	767.0	731.3	367.1	195.0
715.7	110.4	978.7	523.3	213.6	109.2	442.0	294.0	190.1	783.1
941.0	177.1	539.3	412.3	886.2	366.6	206.3	524.7	367.8	187.4
827.2	533.1	923.0	932.7	174.4	314.6	426.2	43.7	861.6	117.3

8.2. Simple Random Sampling

Use this approach when the population can be enumerated as a list (e.g., lab samples, batch numbers, individual organisms in a database) (Krebs, 1999).

1. Define the target population and compile a complete, numbered sampling frame.
2. Determine the required sample size (n) using a power analysis.
3. Use `sample()` in R to draw n unit numbers without replacement.
4. Document the random seed (if reproducibility is desired), the selected unit IDs, and the date of selection.
5. Collect the data or specimens from the selected units.

8.3. Grid-Based Random Sampling (Field)

Use this approach when the sampling universe is a spatially defined area (Elzinga et al., 1998).

8.3.1. Establishing the Baseline Grid

1. Locate the boundaries of the study site.
2. Drive a stake at the origin point and lay two tape measures at 90 degrees to create X and Y axes.
3. Secure the tapes with ground pins to define a coordinate grid (e.g., 20 m $\times$ 20 m).

8.3.2. Generating Coordinates and Placing Quadrats

4. Generate pairs of random integers within the range of each axis:

```
x_max <- 20; y_max <- 20; n_quadrats <- 15
coords <- data.frame(
  X = sample(0:x_max, size = n_quadrats, replace = TRUE),
  Y = sample(0:y_max, size = n_quadrats, replace = TRUE)
)
coords <- coords[order(coords$X), ]
coords$QuadratID <- seq_len(nrow(coords))
knitr::kable(coords, row.names = FALSE)
```

X	Y	QuadratID
0	2	1
1	7	2
2	4	3
4	17	4
7	4	5
7	11	6
7	3	7
8	8	8
8	10	9
9	18	10
13	0	11
17	18	12
17	1	13
18	0	14
18	16	15

5. Walk along the axes to each coordinate intersection and place the bottom-left corner of the quadrat at that point.
6. Never reposition a quadrat after placement; if a coordinate falls on an obstruction, record the obstruction and sample as-is (Elzinga et al., 1998).

8.3.3. Recording Data Within Quadrats

7. Quantify the target variable using one of three standardized metrics:
 - **Density:** Count every individual within the frame
 - **Percent cover:** Estimate the fraction of ground area covered
 - **Frequency:** Score presence (1) or absence (0)
8. Apply the border rule consistently (§4).
9. Aim to sample at least 5–10% of total grid area, or a minimum of 30 quadrat placements, for adequate statistical power (Krebs, 1999).
10. Calculate the mean count (or cover) per quadrat and extrapolate:

$$\text{Estimated population} = \bar{x}_{\text{quadrat}} \times \frac{\text{Total site area}}{\text{Area of one quadrat}}$$

8.4. Stratified Random Sampling

Use this approach when the study area contains distinct environmental gradients that would increase variance under simple random sampling (Elzinga et al., 1998; Krebs, 1999).

1. Divide the study area into homogeneous strata based on a relevant environmental variable.
2. Within each stratum, establish an independent coordinate grid following §8.3.
3. Generate random coordinates separately for each stratum.
4. Sample each stratum independently; do not pool random selections across strata.
5. Weight stratum means by the proportion of total area each stratum represents when calculating site-level estimates.

8.5. Map-Based Sampling with Exclusion Zones

1. Obtain or create a georeferenced map of the study area.
2. Delineate exclusion zones and calculate the remaining accessible area.
3. Generate random coordinates across the full extent, then discard any that fall within exclusion zones.
4. Continue generating coordinates until the target sample size is met within the accessible area.
5. Record discarded coordinates in the sampling log.

R Implementation

An R script for map-based random sampling with exclusion polygons can be built using the `sf` package. Generate random points with `sf::st_sample()`, then filter against exclusion polygons with `sf::st_difference()`.

9. Data Management

1. Record all sampling metadata (date, site, grid dimensions, random seed or table starting position, personnel) on the field data sheet.
2. Enter field data into the project database within 48 hours.
3. Archive original data sheets (scan if paper).
4. Version-control all analysis scripts per [SOP-L801 RStudio and GitHub](#).

10. References

- Krebs (1999)
- Elzinga et al. (1998)
- APHA et al. (2012)

APHA, AWWA, & WEF. (2012). *Standard methods for the examination of water and wastewater* (22nd ed., p. 1360). American Public Health Association.

Elzinga, C. L., Salzer, D. W., & Willoughby, J. W. (1998). *Measuring and monitoring plant populations* (BLM Technical Reference 1730-1). Bureau of Land Management.

Krebs, C. J. (1999). *Ecological methodology* (2nd ed.). Benjamin Cummings.

11. Related SOPs

SOP	Title	Relationship
F001	Field Safety	Required prerequisite
F006	GPS and Site Navigation	Navigate to random points
F603	Vegetation Transect Surveys	Applies this sampling design

12. Training Record

Before designing or executing random sampling independently, personnel must:

1. Read this SOP in full
2. Complete [SOP-F001 Field Safety](#) training
3. Demonstrate proficiency with random number generation in R
4. Participate in a supervised field sampling session using the grid-based procedure
5. Be signed off by the PI or field lead