

T1.1 SAMPLING

What is sampling?

The practical implementation of precision farming requires that the spatial and temporal variability of soils and crops be measured and understood. It is usually insurmountable and expensive to measure everything, but it is also almost always unnecessary to do so. Survey sampling and stereology provide mathematically valid techniques to make reliable inferences about spatial populations based on small samples (Wulfsohn, 2010).

There are two different approaches to statistical inference from a sample:

- **Design-based approaches** – they provide unbiased estimators of the parameters of interest by using a well-defined random selection procedure.
- **Model-based approaches** - the random variation is assumed to arise from intrinsic randomness in the population.

Researchers in precision agriculture and soils have mostly relied on model-based sampling methods. Models of the population can be of considerable practical advantage when based on knowledge of the natural phenomena (e.g. climate, dynamics of weed populations, predators, or disease) influencing the distribution of the population or when models (e.g. regression relations, process-based crop models, soil property functions) are available describing a relationship between the variable of interest with an auxiliary variable. Geostatistical interpolation methods are model-based.

Main sampling types

Systematic random sampling

Systematic sampling has the desired property ensuring geographical spread. Systematic uniform random (SUR) sampling often offers appropriate balance between the estimator precision and the time spent to obtain samples.

Three examples of uniform random (UR) sampling designs for selecting plots in an arbitrarily shaped field, all with the same sampling fraction, are illustrated in Figure 1 (Wulfsohn, 2010).

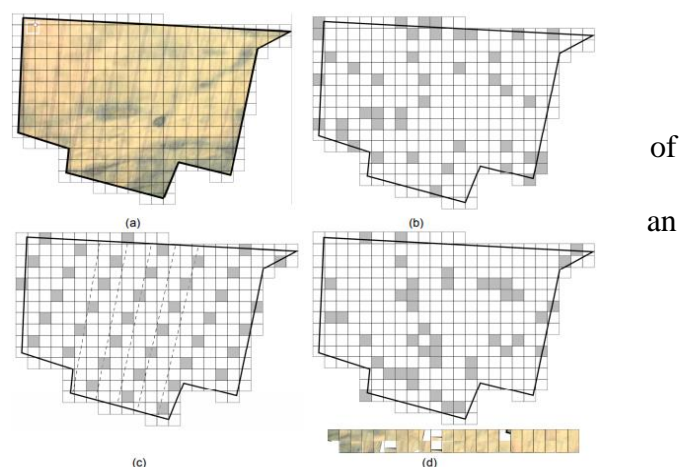


Figure 1: Examples of uniform random (UR) sampling designs for selecting plots in an arbitrarily shaped field
Source: Wulfsohn, 2010

Probability proportional to size sampling

If a measure of size can be attached to each unit, a probability proportional to size (PPS) sample can be drawn (FAO, 2015). According to Wulfsohn (2010), image analysis is used to automatically assign weights to the sampling units; the units are put in a smooth order based on their weights (an optional step), and then a sample of specified size n is sampled systematically on the ordinate of the cumulative weight using a sampling period of Z/n with random start (Figure 2).

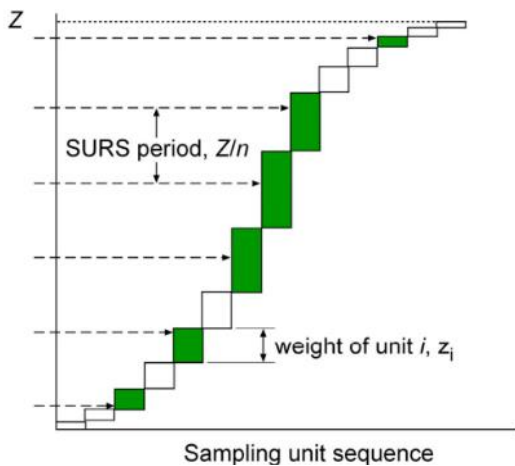


Figure 2: Proportionator (PPS) sampling
Source: Wulfsohn, 2010

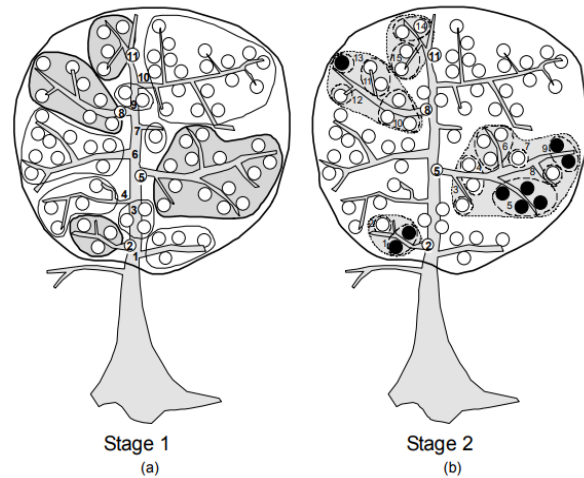


Figure 3: A two-stage fractionator sampling design
Source: Wulfsohn, 2010

Multistage designs

A “level” is a sampling design in which the measurements are made at a constant scale. There are several reasons for using multilevel (multistage, nested) designs. To obtain a sample of seeds from a vegetable crop, it may be convenient to first take a sample of rows from a field, then a sample of fruit from plots located along the sampled rows, and finally take a sample of seeds from the sampled fruit. Indices obtained from remote sensing images or soil maps might be used to design stratified, Smooth, or PPS designs for field and orchard crops (Wulfsohn, 2010). Basso et al. (2001) used remote sensing to target sampling of soil and crop for input to a crop model and combined remote sensing with model output to identify management zones and to interpret yield maps.

Cluster sampling

The main characteristic of cluster sampling is that the sampling unit is a cluster of units. To select a cluster sample, a simple random sample of clusters is taken and each unit in the selected clusters is investigated. Systematic or replicated sampling can also be used to select a cluster. The efficiency of cluster sampling improves when the variability of the sampling units within is large. However, since clusters for agriculture are defined geographically, they tend to be homogeneous. In these cases, more clusters will have to be selected and then subsampled, measures of size. However, cluster sampling in agriculture is a powerful tool – in terms of time and cost – for developing a sampling frame using it as a basis for sample surveys (FAO, 2015). The fractionator design presented in Figure 3 is a two-stage nested systematic cluster sampling design.

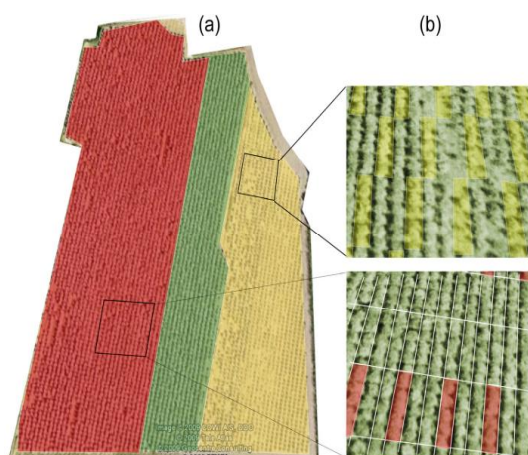


Figure 4: A stratified cluster sampling design
Source: Wulfsohn, 2010

design. The primary sampling units are primary branches, while the higher-order branches serve as second-stage clusters.

Stratified sampling

In stratified sampling, the population is first divided into subgroups called strata, in a so-called stratification process. Then, elements are sampled from each stratum (subgroup) based on a given probability sample design, such as simple random sampling (FAO, 2015). Examples of possible strata are soil horizons, biogeographical areas, morphological zones within plants, and management zones based on easily measured ancillary variables such as vegetation indices, yield productivity, soil properties, and/or past field history. The benefit of stratification is gained by using different sampling fractions in the different strata (Wulfsohn, 2010).

Directed or targeted sampling

Several studies have used ‘directed’ or ‘targeted’ sampling designs, in which sampling is carried out in regions of the field or crop that appear to be visually different from the rest of the field. A frequent motivation is to make accurate prescriptions based on relatively few samples.

Geostatistical sampling

The geostatistical method of local prediction, kriging, depends on having accurate variograms and spatially dependent data from which to predict. All methods of interpolation assume implicitly that data are spatially dependent, which means that sampling should be at an interval that is well within the correlation range of spatial variation. Soil and crop properties can vary at markedly different spatial scales both within and between fields. Therefore, when designing a sampling scheme, it is essential that the spatial scales of variation in the properties of most importance for precision agriculture management are used to guide sampling.

Links to relevant topics

FAO (2015). Handbook on Master Sampling Frames for Agricultural Statistics Frame Development, Sample Design, and Estimation. Retrieved from: <https://www.fao.org/3/ca6398en/ca6398en.pdf>. Accessed on 20.10.2022.

Kerry R., Oliver, M.A., Frogbrook, Z.L., (2010) Sampling in Precision Agriculture. Chapter. July 2010. DOI: 10.1007/978-90-481-9133-8_2. Springer Science+Business Media B.V.

Wulfsohn, D., (2010). Sampling Techniques for Plants and Soil. Landbauforschung Völkenrode, Special Issue 340, 2010.

Key words

design-based approaches

model-based approaches

systematic random sampling

probability proportional to size sampling

multistage designs

cluster sampling

stratified sampling

directed or targeted sampling