SAMPLING ARTHROPODS

Steven E. Naranjo
USDA-ARS, Western Cotton Research Laboratory
Phoenix, Arizona, USA

Sampling is a fundamental component of any experimentally based research program in the discipline of entomology, whether conducted in the laboratory, greenhouse or field. It also is an essential element for describing, measuring and quantifying arthropod population dynamics, whether the goal is to understand the population or community ecology of a given species or group of species, or to develop decision aids for integrated pest management (IPM). Although a population census may be highly desirable for many reasons, rarely does one have the time and resources to exhaustively count every insect, spider or mite within a defined area even if that area is relatively small, say, a small greenhouse or field plot. Instead we must resort to drawing a sample from the population and using the information gained from this sample to estimate variables such as population density or size and to draw inferences that can be applied to the entire population. Fortunately, there is a large body of theory and practical knowledge to guide our efforts in sampling. This short review will attempt to provide a broadly-based, non-mathematical overview of sampling, from discussion of the goals of sampling, to the basic tools for gathering data, to the basic components and mechanics of developing a useful sampling plan.

Samples and sampling are terms with which most entomologists are intuitively familiar. However, as with any field of study, there is a specific vocabulary associated with the study of sampling that may not be as well understood. Thus, brief definitions of some of the more commonly used terms in sampling are provided (Table1).

Goals of sampling

The first, and perhaps most important, element in the development of a useful sampling plan or program is to clearly define the goal(s) of sampling. Is the goal to survey for the abundance or presence of one or more insect species over a large geographical region? Or is the goal to intensively study the temporal dynamics and examine rates of mortality of a single insect species in a relatively small area? Perhaps the goal is to develop a simple procedure that a field scout can use to determine the need for remedial pest control in a field crop. A researcher may be interested in testing the effect of a control tactic on populations of a pest insect using replicated field plots. Although some of the information that will be necessary to devise sampling plans is shared among all these and other applications, as we will see below, there are also differences in the sampling methods that may be used and the levels of accuracy and precision that may be required to provide useful data. For example, a quick and simple relative sampling method such as a sweep net or a colored sticky trap may be sufficient for survey detection or for monitoring the relative abundance of one or more insect species over a large region. Here the premium would generally be placed on extensive spatial coverage at the expense of precise estimation of density, which may not be necessary. In contrast, detailed population studies involving the construction of life tables would require an absolute sampling method, such as quadrat or whole plant counts, and fairly large sample sizes so that mortality factors and recruitment to subsequent life stages could be precisely estimated. The development of a decision-aid for pest scouting could use either an absolute or a relative method of sampling depending on the pest and would likely involve the use of a sequential sampling plan that would allow the classification of pest density as either above or below an established economic threshold.

The time and resources that one has to devote to the activity of sampling are almost always in short supply, even in well-funded research programs. Thus, clear delineation of the goal(s) of sampling is the first step towards ensuring that available time and resources will be put to the best use to answer the questions that are being asked of the sample data.

Sampling methods

Many sampling methods have been devised by entomologists for sampling all types of insects, mites and spiders. Each arthropod poses its own unique challenges. In fact, there are probably as many sampling methods as there are entomologists using them. It would be impractical to try and cover all of the available techniques that have been developed. Instead, focus will be placed on relatively few methods that have been widely used to provide a flavor of the diversity. Readers are referred to the excellent and comprehensive treatment of all aspects of sampling provided by Southwood (1978).

For purposes of organizing this discussion, it is convenient to classify sampling methods according to the type of information that they can potentially provide the sampler. These include absolute methods, relative methods, and population indices. Sometimes the difference between an absolute and a relative method is not distinct and a technique that provides an absolute method for one insect will not do so for another in the same habitat. Absolute sampling methods attempt to provide an estimate of density from a specific and quantifiable fraction of the habitat in which the arthropod lives. Generally, these counts have the unit of numbers per square meter or some other unit of ground area or space within structures. Common examples include quadrat sampling in which a square, rectangle or circle of known area is placed on the ground and all arthropods of interest falling within that area are counted *in situ* or the vegetation is collected and later examined in the laboratory. In agricultural, horticultural or rangeland situations the quadrat delineates a volume that may be relatively small if dealing with low-growing grasses or forbs, or relatively large if dealing with corn plants or fruit trees. Quadrats may also be used to sample arthropods within barns, houses or other structures. Often, the quadrat is placed the day before sampling to minimize disturbance. For arthropods that are found only on the plant surface, absolute counts can be made by counting all individuals on a single plant, groups of plants, or even specific parts of plants such as whole leaves, stems, or branches. Sometimes these efforts are aided by the use of various sorts of **enclosure devices** (e.g., cages or sacks) to confine the arthropods until they can be counted. In many cases, conversion to unit ground area can be accomplished by knowing the number of habitable plants, leaves or branches per unit area. For certain arthropods that are readily dislodged from plants, various kinds of beat cloths and buckets can be used (Fig. 1). Suction samplers may provide absolute counts for certain arthropods depending on plant size and growth characteristics. Absolute sampling of soil-dwelling arthropods can be achieved by taking soil cores or digging trenches of a known volume. Surface-dwelling arthropods can be counted directly or collected from the litter within a quadrat. Arthropods can be extracted from collected soil, plant or litter material using a variety of approaches including simple inspection, dry and wet sieving, flotation in concentrated salt solutions, leaf brushing machines, and other methods. Live arthropods can be extracted by chemical fumigation or heat (e.g., Berlese funnel). Emergence traps can be used in some instances to measure the absolute density of adult stages of insects that pupate in or on the soil. Suction traps or aerial nets towed by an airplane or vehicle can be used to obtain absolute samples of flying insects (Fig. 2). In this instance the unit is a given volume of air that can be calculated by knowing the volume of air drawn by the motor per unit time (or speed of the vehicle) and other environmental factors such as wind speed.

Relative methods are so named because they generally provide estimates of density that are comparable in relation to other estimates made in the same way in the same type of habitat. Relative sampling results in counts per unit of effort and it is not always possible to define the physical units of a method. In contrast with absolute methods, relative methods are generally less costly, easier to perform, and tend to concentrate arthropods. As such they are well-suited to extensive detection and survey work and as components of decision aids in IPM. Relative methods are also widely used in experimental field work where it is often necessary to sample a large number of plots in a short period of time, and only comparative results are sought. In some cases relative estimates can be converted to absolute estimates, but this depends on the specific arthropod and the nature of the sampling method. Common examples of relative methods include the ubiquitous sweep net which is almost the *de facto* standard for estimating densities of arthropods associated with many field crops (e.g., cotton, soybean, small grains, alfalfa). The sweep net can cover a large area of habitat in a short period of time and the particular pattern and number of swings can be standardized across samplers. However, it also samples only a proportion of the population (that inhabiting the tops of plants) and this proportion can change with plant growth, environmental conditions, and time of day. Various methods involve the dislodgement of arthropods from plants into buckets, cloths, nets, trays or other surfaces. Depending on the behavior and biology of the arthropod and the portion of the plant sampled, this may provide either absolute or relative sampling information. In most cases only a small portion of the plant is sampled. The same is true of various suction sampling devices such as the well-known D-vac, modified leaf blowers, or high-powered vacuums attached to carts or tractors. Visual inspections of various plant parts such as leaves or branches are generally considered relative methods, although in many cases, counts on these parts can be related to counts per plant or tree. Timed-counts, which simply involves enumerating all the arthropods that can be counted in a specified period of time are commonly used, especially when the goal it to compare densities of a specific arthropod in a number of different habitats.

By far, traps of various kinds represent the broadest diversity of relative methods. Traps and trapping are the subject of an entry in this volume and so will only be briefly discussed here. Traps can be delineated into those that passively intercept and those that attract actively moving arthropods. Examples of the former include canopy traps, window-pane traps, malaise traps, pitfall traps, and white or clear sticky traps. Examples of the latter include semiochemical- or food-baited traps, traps of various colors (yellow is most common) generally coated with a sticky material, light traps, and sound traps. Mammals and birds also may be used as baits to trap arthropods of medical and veterinary importance. In general, trap counts are very difficult or impossible to convert to absolute counts because they rely on behavior that can be influenced by an array of biological and environmental factors.

Finally, the product or effect of an arthropod population rather than counts of individuals can be used to gauge relative abundance, an approach said to be a population index (plural, indices) because the insects *per se* are not being tabulated, only their activities. Common examples of products include the collection of insect frass underneath plants or trees, the collection of larval or pupal exuviae, or the counting of spider webs. Examples of arthropod effects include the assessment of defoliation by foliage-feeding insects, tunneling by stalk-boring insects, or root pruning soil-dwelling insects (Fig. 3). Acoustic monitoring can be used to detect the presence and relative abundance of insects inside fruit, grain bins or within the walls of structures. Both products and effects can sometimes be related to more quantitative measures of population density. The examination of arthropod effects can be important in their own right in terms of assessing economic damage.

Development of a sampling plan/program

A sampling plan or program is a structured set of rules that guide sampling activities. The sampling plan includes delineation of the sample universe, timing of sampling, the size and nature of the sample unit, how many sample units need to be collected, and how these sample units are spatially allocated among potential sample units in the population. Thus, a sampling plan is distinct from a sampling method or technique (a component of a plan) with which it is sometimes mistakenly confused in the entomological literature. Not all sampling activities in entomology require a sampling plan; for example, collections for systematic or taxonomic purposes. However, if the goal includes the estimation or classification of the density or size of an arthropod population, then the user would be best served by considering, at the very least, selected components of a sample plan to ensure that the sample data are trustworthy. Binns et al. (2000) suggest that a trustworthy sampling plan should adhere to four criteria. First, estimates generated by the plan should be representative of actual pest density thereby avoiding or at least being able to account for any bias. Second, sample estimates should be largely independent of the sampler and of uncontrollable environmental variables such as weather. Third, sample information should be relevant to the questions being asked. For example, estimates of pest density in a management context should be have an identifiable relationship to crop damage. Finally, the sampling plan must be practical for its intended purpose. For example, it would do little good to develop an elaborate and expensive sampling plan for IPM decision-making that scouts would be unable to implement due to time and resource constraints. These criteria can be met by careful attention to the building blocks of a sample plan.

Sample plan components

Delineation of the sample universe is often dictated directly by the goals of the sampling plan. For example, the population of interest for IPM of a particular pest species is generally a single crop field or perhaps a set of crop fields on a single farm that share a similar cultivar and planting date. In a rangeland setting the sample universe might be an entire ranch or perhaps sections or partial sections of the operation. The population of interest in an experimental research program would be a replicate plot representing a given treatment. If the population of interest occurs in an area where physical portions of the habitat are distinct and vary in their suitability for the organism under investigation, several sampling universes might be identified or a stratified sampling approach might be employed. This situation will be discussed below.

Likewise, the timing of sampling is often largely dictated by the goals of sampling. Each arthropod species has a distinct phenology and seasonality that will determine the appropriate time to gather samples on the life stage or stages of interest. The timing of occurrence of particular species might be estimated by phenological models, the presence of indicator plants, or historical records. In instances where the timing of occurrence is not well-characterized or highly variable, the use of detection traps may help to determine the need for more intensive sampling by other methods.

Of all the components of a sampling

plan, the sample unit is perhaps the most important yet most frequently overlooked element. The sample unit defines the unit of habitat represented by the sample and is the foundation for all further development of an efficient sampling plan. Each sample unit should have an equal chance of being selected from the sample universe. It should contain a consistent fraction of the population over time. Thus, definition of the sample unit may include a time component. For example, arthropods may inhabit different parts of a plant over the course of a day due to changes in temperature, light and other factors. The sample unit should be easy to delineate in the sample universe. The sample unit should be appropriately sized relative to the size and behavior of the species in question. For example, a small disk from a leaf may be suitable for spider mites or aphids, but not for large and mobile insects such as assassin bugs or grasshoppers. Finally, the sample unit should strike a balance between variability of counts and the cost of sampling that unit. Using these criteria helps to ensure that the sample unit is representative of the population and help to minimize or eliminate bias.

The sample unit is based on the sampling method from which it is derived. Examples of sample units include 25 sweeps with a sweep net along a single crop row between 0800 and 1000 hours, the lower surface of a leaf five nodes below the terminal apex of a plant, the silks at the tip of one ear of corn, a 8 x 8 cm yellow sticky card oriented horizontally at the top of the crop canopy and exposed for 24 hours, five beats to a tree branch one meter above the ground with a stick over an open 38 cm diameter net, 0.001 cubic meters of soil dug from within a 100 square cm frame, the head or leg of a single cow, or 2 minutes of suction from a D-vac placed over the top third of a crop plant. Very often the sample unit is selected *a priori* based on knowledge of the biology and behavior of the organism and experience, and is not the subject of further investigation. However, careful attention to the size and nature of the sample unit can greatly increase the efficiency of the overall sampling program by minimizing both variability and cost. Established methods are available for comparing various candidate sample units (see Southwood 1978) based on consideration of precision and cost. However, as a guideline, smaller sample units are generally more efficient than larger units. This generalization is based on the typical aggregated dispersion of many arthropod populations and the lower cost of counting individuals on smaller sample units even when a comparatively larger sample size may be required.

Dispersion, or spatial distribution, is a characteristic of populations and can be influenced by a host of biological, behavioral, ecological and environmental factors. Patterns of dispersion are typically categorized as uniform, in which individuals are evenly spaced throughout a habitat, random, in which the location of one individual bears no relation to the location of others, or aggregated (contagious), in which individuals are found in clumps of varying sizes. Many arthropod populations are described as having an aggregated dispersion. Knowledge of dispersion is an important element in the development of sampling plans, influencing both the determination of sample size and the allocation of sample units. The measurement of dispersion is spatially contextual. That is, patterns of dispersion depend on the spatial scale under which they are observed. As a result, the dispersion of a population is greatly influenced by the size of the sample unit. For example a population could be aggregated at the level of a whole plant, but randomly distributed at the level of a single leaf. Thus, in sampling work, dispersion generally refers to the sampling distribution rather than the true spatial distribution. A number of methods are available to explicitly measure dispersion or to provide indices of dispersion. A common index is to simply divide the sample variance by the sample mean. Values <1, ~1 or >1 indicate uniform, random or aggregated distributions, respectively. For sample plan development it is more typical to characterize dispersion with probability models or empirical mean-variance models. Common probability models include the Poisson for random distributions, and the Negative-binomial for aggregated distributions. The primary limitation of using probability models is that sampling distributions typically vary with changing arthropod density and it is not always possible to account for these changes in the model. Empirical regression models offer a solution to this problem by permitting prediction of sample variance from the sample mean. Such models can provide useful information over a wide range of densities. Common examples include Taylor's power law and Iwao's patchiness regression, both of which can be parameterized with simple linear regression.

Once the sample unit has been selected and the sampling distribution has been characterized by either a probability or empirical model, the sample size can be calculated. How this is done depends on the goals of sampling, the criteria for determining the adequacy of the population estimate or classification, and the resources that can be devoted to sampling activities.

Types of sampling plans

Fixed-sample-size sampling plans, as the name implies, are based on the collection of a set number of sample units. The optimum sample size is the minimum number of sample units that achieves a desired level of performance. Fixed-sample-size sampling can be applied to density classification as well as estimation; however, it is more commonly used for the latter and so consideration will be restricted to estimation here. Several approaches can be used depending on how dispersion is characterized and how precision is quantified. The simplest approach is to calculate the optimum sample size such that the standard error is within a specified proportion of the mean, or that the mean falls within a certain confidence interval of the true mean with a specified probability. Variations on this approach explicitly account for sampling distribution (e.g., Poisson, Negative binomial). A slightly more complicated and robust approach uses a mean-variance model such as Taylor's power law to solve for sample size using the same measures of precision. The main advantage of this approach is that the optimum sample size is dependent on the mean density of the population. Typically, sample size decreases exponentially with increasing density. The reader is referred to Southwood (1978) and Pedigo and Buntin (1994) for mathematical details.

One of the main advantages of a fixed sample size is that sample allocation (see below) can be structured so that sample units are collected from throughout the sample universe. The main disadvantage is that the mean density is unknown for the population under study at any given point in time. One solution would be to calculate sample size for the lowest expected density or the average expected density and then apply that in all cases. This would clearly be inefficient at high densities, or require too many or too few sample units depending on density at any point in time, respectively. The use of double sampling would be an alternative solution. Here a small sample would be collected and used to solve for the required sample size at that point in time.

Sequential sampling is another alternative that solves the problem of unknown density and generally maximizes sampling efficiency. Sequential sampling was developed for military application during World War II and has since become the most widely applied approach for developing sampling plans for arthropods based on both density estimation and density classification. Efficiency is optimized because in sequential sampling the need for further sample information is assessed following the collection of each individual sample unit. Regardless of the mean density, sequential sampling ensures that no more sample units are collected than necessary in order to achieve a predetermined level of precision or classification accuracy (Fig. 4). Sequential sampling for estimation of mean density operates by accumulating counts over subsequent sample units and then consulting a ‘stop-line’. This stop-line represents the cumulative count as a function of sample size and the desired precision. It can be plotted or presented in table form. Matching the current cumulative count and sample size relative to the stop-line determines whether more sample units are need or whether sampling should be terminated. Once sampling is terminated the mean is calculated by simply dividing the cumulative count by the number of sample units collected. Sequential sampling plans for mean estimation can be developed on the basis of probability or mean-variance models. Frequently, a minimum and maximum sample size is specified as part of the sequential plan. The reader is referred to Nyrop and Binns (1991) and Pedigo and Buntin (1994) for mathematical details.

A second application of sequential sampling involves the classification of population density. Most often this approach is applied to the development of decision-aids for IPM where one simply wants to know whether pest density is above or below the economic threshold.

Operationally, sequential sampling for classification is similar to that described above for mean estimation except that a critical density (e.g., economic threshold) and upper and lower decision boundaries must be specified. The calculation of the stop-lines is also more complex. Several approaches have been developed. One goes by the daunting name of the sequential probability ratio test (abbreviated SPRT) and can be based on various probability models including the Poisson, negative-binomial, normal or binomial. A second technique termed Iwao’s confidence interval method is based on the normal distribution, and a third called the sequential interval procedure is based on the binomial distribution. All three approaches result in pairs of stop lines rather than a single one, delineating three distinct decision zones. More complex plans can be developed involving multiple stop lines designating even finer decision zones. Such sampling plans are popular for use in IPM because of their high degree of efficiency. Very few sample units are required when densities are much smaller or larger than the critical density and the maximum number of sample units is required only when the density is very near the critical density. Sequential sampling plans for classification are developed and evaluated on the basis of two criteria, the operating characteristic (OC) and the average sample number (ASN). The OC estimates the probability of classifying density below the critical density as a function of the true mean density. Various parameters of the sequential plan can be adjusted to increase or decrease the level of classification accuracy and the associated sample size.

Up to this point the discussion has focused on enumerative sampling in which all individuals inhabiting a sample unit are counted. However, depending on program goals, the general efficiency of sampling can be potentially improved by determining only the presence or absence of individuals in the sample unit. This approach, referred to as binomial sampling, has been widely used in the development of decision aids for IPM and is typically implemented as a sequential classification sampling plan as described above. Generally, the foundation for this approach is the development of a relationship between mean density and the proportion of sample units containing at least T individuals. T is known as the tally threshold and traditionally has a value of 1 (true presence/absence); however, it can take on any value and proper selection of T can often improve the accuracy of classification. This mean density-presence/absence relationship can be determined using various probability models or it can be represented by one of several empirical models. In some instances it is more convenient to base the economic threshold simply on binomial counts and, thus, there is no need to develop a relationship with mean density. A well-devised and tested binomial sampling plan can significantly reduce the cost of sampling without sacrificing accuracy of density classification in IPM decision-making. Binomial sampling can also be an efficient approach for estimating population density, although its application for this purpose has been limited. The reader is referred to Nyrop and Binns (1991), Pedigo and Buntin (1994), and Binns et al. (2000) for mathematical details on sequential sampling for classification using both enumerative and binomial sampling.

A final sampling approach, termed variable-intensity sampling, is a classification method that shares some of the attributes of both fixed-sample-size and sequential sampling plans. It was developed primarily to address a potential deficiency in sequential sampling having to do with adequate coverage of the sample universe. As a result of its high efficiency, sequential sampling often terminates sampling after only a small portion of the sampling universe is observed. Variable intensity sampling solves this problem by delineating a transect (see allocation below) across the sample universe, dividing that transect into equal lengths, and requiring at least one sample unit in each section. Thus, the sample is more representative, especially in situations where pest density might be highly variable across a field. The overall number of sample units taken in each transect section ultimately depends on prior sampling information much as in a traditional sequential sampling plan. Variable intensity sampling can be based on enumerative or binomial counts.

Allocation of sampling units

The final element of a sampling plan is the collection of sample units from the sample universe. Random sampling assumes that every possible sample unit within the sample universe has an equal chance of being selected. As such, random sampling ensures that the sample will provide unbiased estimates (assuming that the proper sample unit is used and sample size is adequate). Random sampling can be accomplished by enumerating every sample unit and then selecting them using a random number table, drawing numbers from a hat, or using a computer random-number generator. This assumes that the spatial location of each sample unit can be correctly identified which may be extremely difficult. In field crops, orchards or greenhouses with evenly spaced plants, sample unit location can be simplified somewhat by defining a coordinate system and randomly selecting a random row and a random number of steps or plants along the row from the edge. A variation of random sampling is **s**tratified sampling. This approach may be used if there is some a priori knowledge that the population being sampled is heterogeneously distributed. Common examples are the differential distribution of insects in border rows compared with interior rows of a crop field, the differential distribution of insects within the vertical strata of a plant, or differences in distributions of arthropods within structures. In instances such as these a more precise estimate may be achieved by allocating sample units to different strata in proportion to the size of each strata, or in relation to variability within each strata. More sample units would be collected in strata that are known to be more variable and vice versa. The allocation of sample units within a stratum is random, preserving the benefits of random sampling, but different formulae are needed to estimate mean density and variance.

Despite the desirable qualities of random sampling, it is rarely used in practice mainly because proper implementation is costly. It is much more common for sampling plans to be implemented using some sort of systematic sampling. In systematic sampling, some predetermined path is chosen and sample units are collected at fixed intervals along this path. Typical examples are diagonal transects, or X-, V-, or Z-shaped patterns laid out across the sample universe. Aside from being easy and inexpensive to implement, systematic sampling generally allows sample units to be collected from throughout the entire sample universe, resulting in a more representative sample. A degree of randomness can be added by randomly selecting the starting point of the transect and the interval between sample units. A disadvantage to systematic sampling is that it can produce biased estimates if there is some underlying pattern in the population that closely mimics the pattern of sample unit collection. Further, there are no exact formulae for calculation of variance and so estimates based on random sampling are often assumed. The overall effect of these limitations is not well-known, however, proper testing and evaluation (see below) can help to ensure that the sample obtained is reliable and provides the necessary levels of precision and accuracy.

Testing and evaluation of sample plans

Ideally, sampling plans should be developed from observations that encompass the geographic area and range of environmental conditions that future users of the sampling plan are likely to encounter. In reality, however, sampling plans are often developed from a more restricted range of observations and then applied to similar or novel situations. In addition, regardless of the extent of data collection, any sampling plan is based on observations that are measured with some amount of error. Thus, it is important that the performance of a sampling plan be evaluated so that its limitations and strengths can be better defined. This evaluation process is known as validation. In the past, generally little attention was paid to sample plan validation, however, this situation is changing and several validation approaches have been formalized in the last 10-15 years. The simplest involves the use of Monte Carlo simulation. In this method the sampling plan is applied to sample units counts that are randomly drawn on a computer from a probability distribution (e.g., Poisson, Normal, Negative binomial) that is representative of the sampling distribution of the arthropod in question. This procedure is repeated a large number of times (usually 500 or more) to represent a large number of possible sampling outcomes and average precision or classification accuracy (and their variances) are calculated. An alternative approach uses real sample data independently collected at the time of sample plan development or collected from an area where the sample plan is to be implemented. These real data are then resampled on a computer much as in the Monte Carlo approach. The advantage of the resampling method is that the data rather than a probability distribution are used to represent the sampling distribution. The result is a more robust test of the sampling plan and the assumed probability model upon which it may be based. A disadvantage is that larger amounts of data collection are required. With both approaches, the validation data can have an explicit spatial arrangement so that issues of sample unit allocation can be tested as well. The reader is referred to Naranjo and Hutchison (1997) and Binns et al. (2000) for further detail and available computer software.

References

Binns, M.R., J.P. Nyrop, and W. van der Werf. 2000. Sampling and monitoring in crop protection: the theoretical basis for developing practical decision guides. CABI Publishing, Wallingford, United Kingdom.
Cochran, W.G. 1977. Sampling techniques (3rd ed.). John Wiley and Sons, New York, New York.

Flint, M.L., and P. Gouveia. 2001. IPM in practice: principles and methods of integrated pest management. University of California, Statewide Integrated Pest Management Project, Agriculture and Natural Resources Publication 3418.

Morris, R.F. 1960. Sampling insect populations. Annual Review of Entomology 5: 243-264.

Naranjo, S.E., and W.D. Hutchison. 1997. Validation of arthropod sampling plans using a resampling approach: software and analysis. American Entomologist 43: 48-57.

Nyrop, J.P., and M. Binns. 1991. Quantitative methods for designing and analyzing sampling programs for use in pest management. Pp. 67-132 in D. Pimentel (ed.), Integrated pest management. CRC Press, Boca Raton, Florida.

Pedigo, L.P., and G.D. Buntin (eds.). 1994. Handbook of sampling methods for arthropods in agriculture. CRC Press, Boca Raton, Florida.

Southwood, T.R.E. 1978. Ecological methods: with particular reference to the study of insect populations (2nd ed.). Chapman and Hall, London, United Kingdom.

Table 1. Definitions of some common terms associated with arthropod sampling.

Absolute method	A sampling method that allows the estimation of numbers per unit of habitat, generally per unit of ground area.
Accuracy	A measure of the closeness of an estimate to the true mean or variance of a population.
Bias	An unidirectional deviation of an estimate from the true mean or variance of a population.
Binomial sampling (presence/absence sampling)	Sampling based on determining the presence or absence of one or more individuals in the sample unit in lieu of complete counting of all organisms. Commonly used for pest management decision-making application. Contrast with enumeration sampling.
Census	Complete enumeration of every individual within a defined sample universe. Contrast with sample.
Classification	A sampling plan that classifies population density as being either above or below some predetermined level (e.g., economic threshold), or belonging within some density class (e.g., low, medium, high). Commonly used in pest management decision making application. Contrast with estimation.
Dispersion/distribution	The spatial patterning of individuals in a population in their habitat. This pattern can be broadly described as uniform, random or, most commonly, aggregated. Dispersion can be quantified by explicit spatial indices, or more typically by probability models (e.g., Poisson, Negative-binomial) or empirical models (e.g., Taylor's power law). The measurement of dispersion is dependent on the sample unit size and population density. Most often the term is used to denote the sampling dispersion/distribution and not true spatial patterning. Dispersion is an important element in the development of a sampling plan.
Double sampling	A sampling approach in which an initial sample (usually small) is drawn and used to determined the necessary sample size for a subsequent sample within the same time period.
Efficiency	A measure of the level of precision or accuracy per unit of cost (time or currency).
Enumeration sampling	Sampling based on the complete counting of all individuals in the sample unit. Contrast with binomial sampling.

Estimation	A sampling plan that numerically estimates population density or intensity. Commonly used for detailed population dynamic and experimental studies. Contrast with classification. The term can also denote the process of calculating various statistical parameters such as variance.
Fixed-sample size sampling plan	A sampling plan in which a predetermined number of sample units is collected based on a prescribed level of desired precision, classification accuracy, or level of confidence. Contrast with sequential sampling.
Mean-variance model	A model, generally a regression model, which predicts the variance of a sample from an estimate of the sample mean. Such models can characterize dispersion over a large range of population densities and are commonly used in developing sampling plans. Common examples are Taylor's power law and Iwao's patchiness regression.
Operating characteristic	A measure of the accuracy of a classification sample. An operating characteristic curve shows the probability of classifying density below a critical value (e.g., economic threshold) as a function of true mean density. Ideally, probabilities should be near 1 at densities far below the critical value, near 0.5 at densities very near the critical value and near 0 at densities far above the critical value. Often abbreviated as OC.
Population index	A sampling method that attempts to provide an indirect estimate of population density based on an associated product (e.g., frass) or effect (e.g., defoliation).
Precision	A statistical measure of the repeatability of an estimate relative to a group of estimates from the same population at the same time. Typically measured as the quotient of the standard error of the mean over the mean. Low numerical values indicate high precision, high numerical values indicate low precision. Precision is a key element in developing and evaluating the performance of sampling plans.
Probability model	A mathematical description of the dispersion or distribution of individuals in a population based on numbers per sample unit. Common models include Poisson, Negative-binomial, Binomial, Normal, and Neyman Type A. Such models can form the foundation of a sampling plan.
Random sampling	A method of allocating sampling units within a sampling universe in which each sample unit has an equal chance of being selected. Random sampling ensures an unbiased estimate, but is rarely used in the practice of arthropod sampling due to time and cost constraints.
Relative method	A sampling method that results in numbers per unit effort. Generally much faster and easier than absolute methods. Sometimes possible to convert counts to absolute density. Common examples are sweep nets and sticky traps.
Robust	Within statistics and sampling, the quality of being widely applicable. For example, a sampling plan is robust if it can be used to precisely estimate the density of an arthropod under many different environmental conditions.
Sample	A collection of sample units from a sampling universe. Contrast with census.
Sampling method/technique	A particular tool or technique used to gather information on population density from the sample universe. Examples include sweep nets, beat clothes, visual counts, suction devices, and various kinds of attractive and passive traps.
Sampling plan/program	A structured set of rules for collecting a sample that is based on delineation of the sample universe, knowledge of dispersion, a specific sample unit, a predetermined sample size (but see sequential sampling), time of sampling, and a given allocation of sample units throughout the sample universe (e.g., random, stratified, systematic).

Sample size	The number of sample units collected from the sample universe for a given sampling effort. Ideally, sample size is based on prior information about dispersion and the desired level of precision or accuracy. Sometimes referred to as average sample size, abbreviated ASN.
Sample unit	A proportion of the habitable sample universe from which counts of individuals are taken. The sample unit is based on a specific sampling method, should be representative of the behavior and size of the organism, and should strike a balance between cost and variability. Examples include whole leaves, branches or plants, a set number of sweeps along a row, a sticky trap set at canopy height, 10-minute timed count, a 1 m2 quadrat, etc.
Sample universe	The physical area that contains the population of interest. Typical examples include a single crop field, an orchard, a section of stream, the cows in a pasture or barn, etc. Synonymous with the management unit in agricultural production systems.
Sequential sampling plan	A sampling plan in which "stop-lines" continually assess the need for additional sample units during the sampling effort on a given day or time interval. These stop lines are typically based on a prescribed level of desired precision or classification accuracy. Sequential sampling is efficient because no prior knowledge of density is required and no more sample units than necessary are collected. Contrast with fixed-sample size sampling.
Stratified sampling	A method of allocating sampling units in which the sample universe is subdivided into 2 or more sections and sample units are randomly collected from each subdivision in proportion to their size or relative variability. Examples include the subdivision of a crop field into border and interior sections, the subdivision of a plant along vertical strata, or different rooms within a structure.
Systematic sampling	A method of allocating sample units within a sample universe in which sample units are collected at fixed intervals along a predetermined pattern. Examples include sampling along transects or along V or X shaped patterns in crop fields. Often, the starting point of the pattern is determined randomly. Systematic sampling is probably the most common method for allocating sample units in most agricultural systems because it is simple and time-efficient.
Tally threshold	In binomial sampling, the number of individuals required to be present per sample unit to consider the sample unit infested. Proper selection of the tally threshold can improve the accuracy of classification sampling.
Variable-intensity sampling plan	A sampling plan that shares the characteristics of fixed-sample size and sequential sampling plans in which prior sampling information is used to evaluate the number and allocation of subsequent sample units, but ensures that sample units are collected throughout the sample universe.

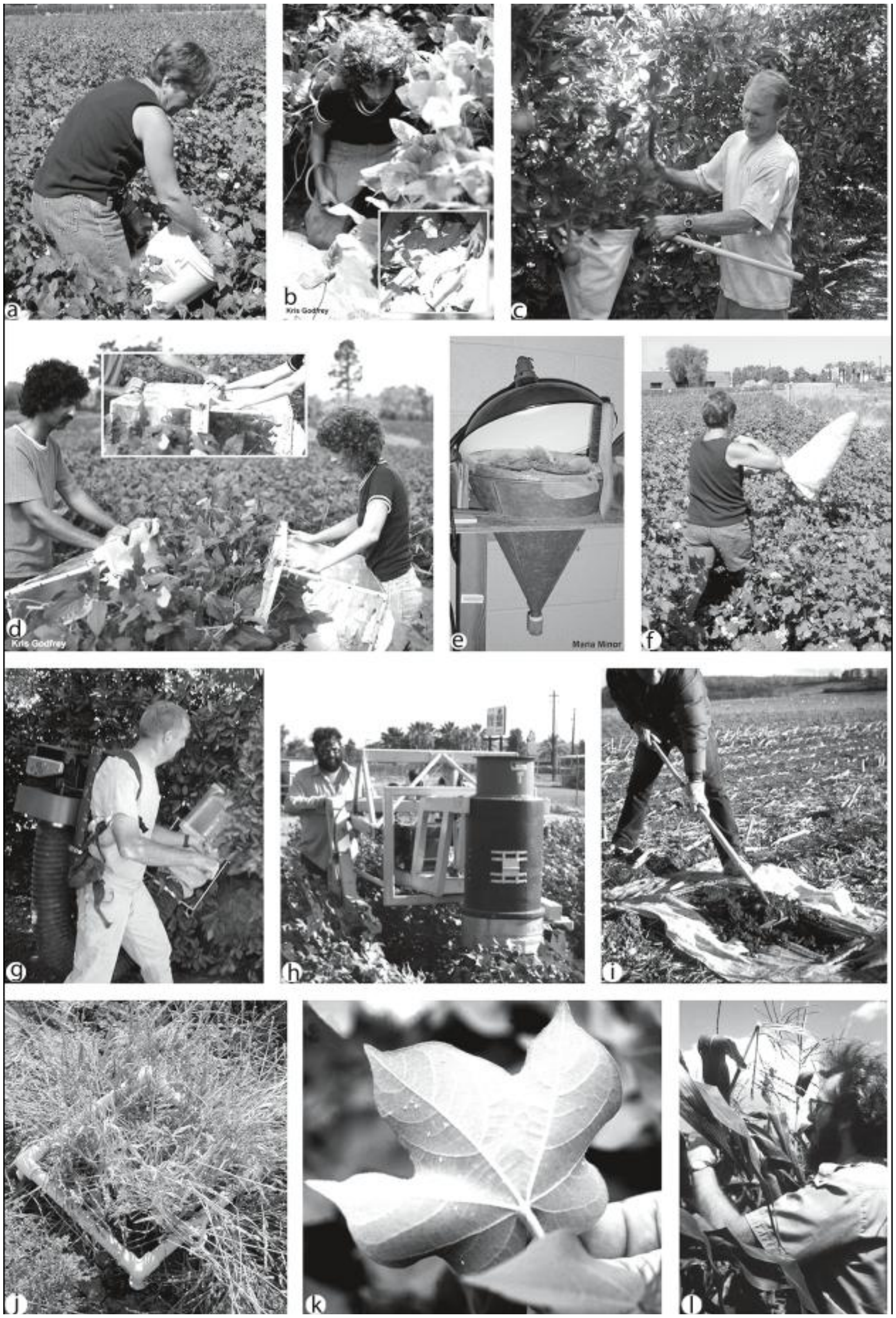

Figure 1. Examples of various kinds of sampling methods: (a) beat-bucket in which a plant is shaken against the sides of a bucket to dislodge arthropods; (b) beat-cloth in which a piece of white cloth is placed on the ground and plants are shaken or beaten to dislodge arthropods; (c) beat-net (or cloth) in which arthropods are dislodged into a net with a beating stick; (d) plant cage that encloses a specific portion of habitat; arthropods are dislodged from the enclosed plant material; (e) example of a Berlese funnel in which arthropods are driven from litter or soil into a collection bucket by the application of heat and light from above; (f) sweep net; (g) D-vac suction sampler; (h) high powered suction sampler designed for field row crops; (i) trench for sampling soil-dwelling arthropods; (j) quadrat; (k) visual inspection of a leaf; (l) visual inspection of a whole plant. All photographs by the author unless otherwise noted.

David Voegtlin

Figure 2. Examples of various kinds of traps for sampling arthropods: (a) canopy trap for sampling insects flying vertically out of a crop field; (b) emergence trap for sampling adult stages of insects that pupate in the soil; (c) window-pane trap for assessing the movement patterns of flying insects; (d) modified malaise trap for assessing the movement patterns of flying insects; (e) one of a large number of sizes and types of sticky traps; (f) suction trap for aerial sampling of flying insects. All photographs by the author unless otherwise noted.

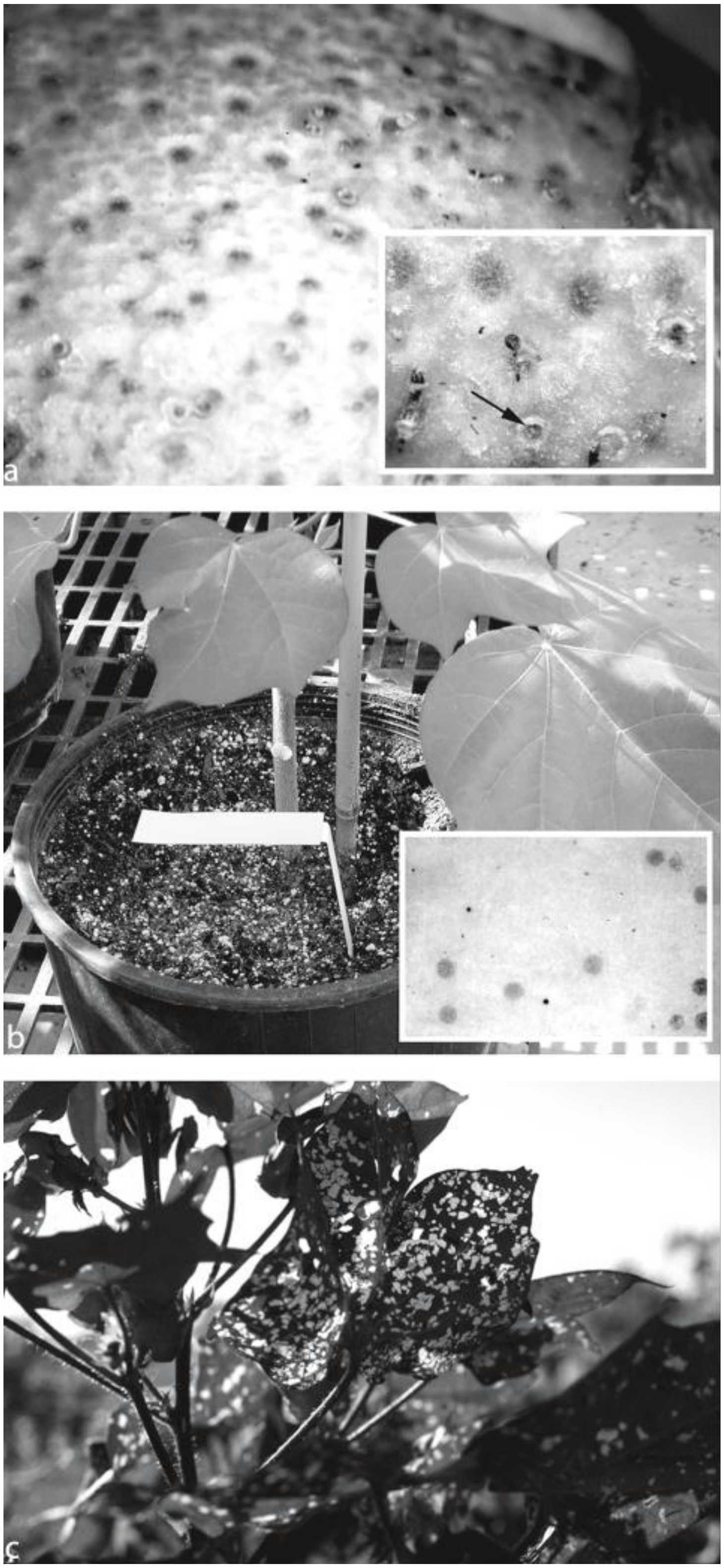

Figure 3. Examples of methods for estimating population indices: (a) counting entrance holes of neonate pink bollworm larvae on the surface of a cotton boll; (b) counting honeydew droplets from whiteflies or aphids on water-sensitive paper; (c) assessing caterpillar abundance from defoliation. All photographs by the author.

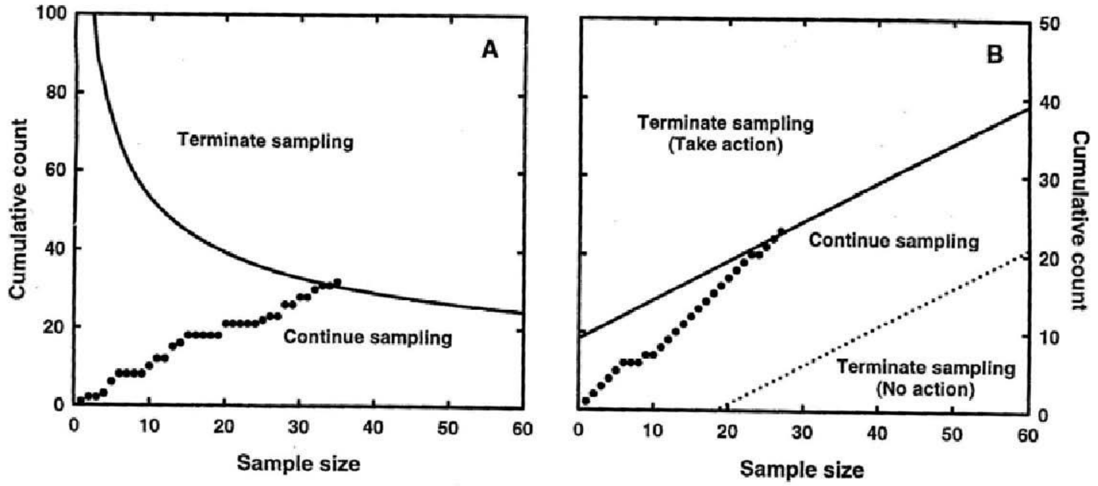

Figure 4. Hypothetical examples of (A) sequential sampling plan for estimating mean density with a fixed precision and (B) sequential sampling plan for classifying mean density relative to an economic threshold. The lines denote the stop-lines and the data points represent a hypothetical sample that terminates after crossing the stop-line.