Society for Range Management

Q1

Sampling design workflows and tools to support adaptive monitoring and management

By Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl and Travis W. Nauman

On the Ground

- Adaptive land management requires monitoring of resource conditions, which requires choices about where and when to monitor a landscape.
- Designing a sampling design for a monitoring program can be broken down in to eight steps: identifying questions, defining objectives, selecting reporting units, deciding data collection methods, defining the sample frame, selecting an appropriate design type, deciding stratification and allocation, and identifying the required sampling effort.
- Here, we provide descriptions of each step in the process and identify tools and resources to complete each step.

Keywords: monitoring, sample design, monitoring workflow, data collection, stratification, spatially balanced sampling.

Rangelands 000():1–9
doi 10.1016/j.rala.2021.08.005

Q2

Introduction

Adaptive land management requires observations of resource conditions, which inevitably entails choices about where and when to measure (i.e., sample) a landscape. This often takes the form of monitoring the collection of data to describe the current state of resources in a landscape.[1] The scope and scale of sampling for monitoring efforts vary dramatically depending on the intended use of the data (i.e., monitoring objectives), from national-scale rangeland health trend monitoring such as the US Department of Agriculture's National Resource Inventory (NRI) [2,3] to evaluating a single management activity such as sampling reference areas for the purpose of evaluating reclamation of an oil well pad.[4,5] Regardless of the particular context, all sampling efforts require a sampling design specifically tailored to the monitoring objectives at hand to effectively define a set of rules describing 1) where sampling might occur, 2) where sampling will occur, and 3) how and when to collect data. Further decisions about data collection methods, sampling design type, stratification, and sampling effort (i.e., number of sampling locations) also require careful consideration of monitoring objectives before finally drawing a sample and beginning a monitoring campaign. The process of creating a sampling design can be difficult, and making poor decisions about sampling design runs the risk of producing uninformative and even unusable data that may lack appropriate inference or result in misleading conclusions. But for each step of the design development process there are tools (i.e., software, benchmarking frameworks, reference manuals, and datasets) available to make informed decisions.

Many of the commonly used resources for developing rangeland monitoring sampling designs[1,6–8] are either focused on narrow applications, such as single objectives at local scales, or do not include recent sampling approaches, such as spatially balanced sampling or combining data from multiple sampling efforts. Accordingly, a need exists to revisit sampling design principles for modern rangeland monitoring applications. Below we present a general workflow of the process for creating a sampling design and provide examples of tools available for each step, using greater sage-grouse (*Centrocercus urophasianus*) habitat monitoring as an example (Table 1). Regardless of the specific steps or order, developing a sampling design is not a strictly linear, one-way process—iteration, reevaluation of decisions, feedback, and documentation are critical. In adaptive monitoring, that iteration will potentially happen after the initial sampling design has been undertaken, as understanding the effectiveness of the design depends on the data produced (McCord and Pilliod, this issue). However, the order presented below can make the process more tractable and smoother.

Table 1
A sample design process using greater sage-grouse (*Centrocercus urophasianus*) habitat as an example

	Process step	Decision points	Tools	Conclusion and rationale
1.	**Questions**	What are the broad questions to answer? What are the goals?	Policy	**There must be acceptable sage-grouse habitat in high priority areas**
2.	**Objectives**	What needs to be known to answer the questions? What are the specific objectives?	Benchmark Tool; technical reference (*Sage-grouse Habitat Assessment Framework*)	**At least 75% of sage-grouse habitat must have between 15% and 25% cover from sagebrush with 80% confidence**
3.	**Reporting Units**	What are the areas or timeframes that data should be summarized over?	Policy	Reporting will be done in **priority areas** for **nesting habitat, brood-rearing habitat, summer habitat, and winter habitat** because each has distinct management needs
4.	**Data Collection Methods**	Given what information is needed, which methods are most appropriate to collect the data?	Technical reference (*Monitoring Manual for Grassland, Shrubland, and Savannah Ecosystems)*	**Line-point intercept** is a common, standardized way to collect cover data and has a well-documented protocol
5.	**Sample Frame**	Given where the project applies to, specifically what parts of the landscape should be sampled to evaluate the objectives?	Benchmark Tool; policy	**Sage-grouse habitat within the priority areas on BLM land**, as required by policy and to include all reporting units
6.	**Design Type**	Given the goals of the project, how should the sample locations be selected from the sample frame?	Technical reference	At this scale, **spatially balanced, probabilistic sampling** should characterize the landscape well
7.	**Stratification & Allocation**	Given the sample frame and reporting units, are there additional measures to take to allocate samples?	Policy; Balanced Design Tool	Because policy requires reporting by priority areas, it makes sense to **stratify by priority areas** to ensure that each is sampled adequately
8.	**Required Effort**	How many samples across what timeframe is enough to meet the goals of the project?	Technical reference; Balanced Design Tool; Benchmark Tool	Based on the expected variability of the cover indicator and the requirement of 80% confidence, **100 sampling locations** should be adequate

Note: Each step includes what questions need to be answered, tools used, and outcomes with a justification. Although the numbering indicates a linear flow, revisiting earlier steps in light of later decisions is common.

The process

Questions

The very first step in creating a sampling design is to establish what questions, goals, or information requirements are being addressed with the planned data collection (Table 1, #1). For some efforts, these may already exist, as with the standards laid out in the Clean Water Act or the Sage-Grouse Habitat Assessment Framework (HAF)[9] in which the questions are laid out precisely. For example, in the HAF, the formal monitoring objectives for the question of "is this suitable sage-grouse habitat?" are determined according to a form that lists the data required and interpretations of the ranges of values to reach a conclusion.[9] In many cases, however, the process is less well defined and can make specifying formal monitoring objectives challenging. For example, a common question such as "Were habitat improvement treatments effective?" requires a number of subsequent clarifying questions to develop monitoring objectives, including "what kind of habitat," "what indicators define that habitat," and "what constitutes effective?" Note the need for clarifying questions is not a weakness or a sign of poorly written management questions, but rather a part of the process.

Objectives

Without clear and explicit articulation of monitoring objectives about what needs to be known, there is no guarantee that a sampling design will produce data that can address the question, goal, or requirement that initially spurred the sampling effort.[10] A clear monitoring objective for the goal to evaluate the effectiveness of habitat improvement treatments might be, "At least 75% of greater sage-grouse habitat in treatment areas must have between 15% and 25% foliar cover from sagebrush." This clearly states what is meant by "effective" and how that effectiveness is being measured as part of the monitoring effort. Specific monitoring objectives may also be subject to important contextual factors that can limit what is possible (e.g., site ecological potential), species-specific habitat requirements (e.g., breeding versus brood-rearing), and potential risk factors (e.g., large amounts of bare ground and risk of accelerated soil erosion). Ideally, monitoring objectives are supported by reference materials like ecological site descriptions and scientific literature as well as professional experience and judgment. Often, specifying quantitative monitoring objectives is the most challenging aspect of sampling design, especially given what is achievable with available resources like staffing and funding and the timeframe for sampling.

Given that most monitoring efforts are limited by available resources, monitoring data are increasingly collected to address multiple resource questions or management goals. Creating specific monitoring objectives is especially challenging and requires even more thoughtful consideration throughout the sampling design and monitoring implementation process if the intent is to use the data for multiple questions or goals. For example, although broad-scale monitoring projects like NRI have primary goals of monitoring overall rangeland health and trends, the data can have valuable applications for other purposes such as species distribution modeling or re-

Please cite this article as: Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl et al., Sampling design workflows and tools to support adaptive monitoring and management, Rangelands, https://doi.org/10.1016/j.rala.2021.08.005

mote sensing applications. Another example is the use of Bureau of Land Management's (BLM) Assessment, Inventory, and Monitoring strategy (AIM) under which data are often collected for Land Use Plan evaluation and also used for more local applications like grazing permit assessments for allotments.[11] Thus, considering secondary uses of data during the design stage can extend the utility of a sampling design for a broader set of end users and purposes.

The framing of monitoring objectives in a quantitative format can help with creating benchmarks for evaluating monitoring data. Benchmarks are defined categories such as "meeting criteria" and "not meeting criteria," which are related to management decisions and can be applied to data to make comparisons between different sampling locations.[12,13] For example, if the management objectives for percent of foliar cover account for ecological potential, then "meeting criteria" for arid shrubland would be a lower percent foliar cover than in a wetter and more productive riparian area. By applying benchmarks to separate values from sampling locations into "yes/no" categories, they are harmonized and can be directly compared across lands with differing ecological potential. The process of creating benchmarks can also be useful for identifying priorities both for sampling design and broader management objectives. Patterns related to data needs and distribution of resources may emerge to guide allocation of sampling effort to meet project goals. For example, if when determining objectives for sage-grouse monitoring, a significant number of suitable habitat criteria relate to breeding habitat, in particular, then sampling design steps focused on breeding habitat areas and seasons would be appropriate while placing less emphasis on other seasonal habitats.

Reporting units

This stage of sampling design is for determining at what level (i.e., what areas or span of time) results should be summarized, commonly referred to as reporting units (Table 1, #3). The levels do not need to be spatially exclusive and can be based on any relevant information including administrative boundaries, soil maps, ecological potential units, management history, or other management-relevant units. Reporting units differ from sampling strata (see "Stratification and Allocation" below) in that they are summary units and not used for dividing an area into regions for sampling efficiency. During the design process, reporting units do not need to be exhaustive, and additional reporting units can be identified and potentially reported on after the design is completed (provided sufficient samples exist in each reporting unit). However, it is helpful to identify all possible reporting units during the design phase to build the design to ensure sufficient sampling effort in each reporting unit. For sage-grouse habitat, knowing results will be needed for each watershed means the sampling design must contain adequate sampling in each watershed, which might not be the case if reporting units were not explicitly considered.

Tracking information about monitoring objectives and reporting units is an important but complicated task, and benchmarking frameworks like the BLM's AIM Terrestrial and AIM Lotic Benchmark Tools (AIMBT)[14,15] can provide valuable structure for a project (Fig. 1). The AIMBT consists of Microsoft Excel workbooks using a standardized format for the description of benchmarks including the applicable data, the qualifying range of values, and the associated reporting units. This structure serves as documentation for later reference and guides the building of appropriate benchmarks. The AIMBT is machine-readable and supported by software tools (i.e., the AIM.analysis R package[16]), which makes applying benchmarks to data at a later time as part of the data analysis easier to automate.

The AIMBT can and should be used in conjunction with other decision tools including policy or technical references directed at monitored resources. For example, in sage-grouse habitat monitoring, many of the objectives and benchmarks are already defined in the HAF. Not all resources have existing quantitative policy guides or requirements, so it also can be useful to reference Ecological Site Descriptions (ESDs) through the Ecosystem Dynamics Interpretive Tool (EDIT), a web interface to a system designed to provide characterization of the ecological potential and behavior of a site using state-and-transition models, and the supporting resource management knowledge[17] (Fig. 1).

Data collection methods

Selecting data collection methods is determined by monitoring objectives and is an important aspect of sampling design (Table 1, #4). Many methods are used for collecting environmental data, but likely only a narrow set will effectively answer any given question or satisfy an objective. Major factors to consider include whether the objectives of the sampling design include requirements for data that are quantitative, qualitative, or both and the required level of precision.[18] Methods differ in sensitivity to variability, and the selection of a method influences the number of sampling locations (i.e., the level of sampling effort or sample size) required to produce useful data. For example, measurements of gaps in foliar cover described in the Monitoring Manual for Grassland, Shrubland, and Savanna Ecosystems (MMGSSE)[1] has been shown to have higher interobserver variability than point-intercept estimates of foliar cover in comparisons made using data available in the BLM Terrestrial AIM Database.[19] The higher observer variability of the canopy gap results may require either a reduced level of acceptable precision or a greater number of sampling locations to produce a high-confidence estimate than for point-intercept estimated foliar cover. If the data will be used in conjunction with other data sets it is important that the methods selected measure the same or compatible variables. Regardless of the requirements, there are certainly documented and robust methods of gathering the needed monitoring data, and a best practice is to use existing documented methods from peer-reviewed sources, which are frequently organized in manuals and technical references[1,18] (Fig. 1). Standardized methods also make comparisons to other published or publicly available data easy, which can help contextu-

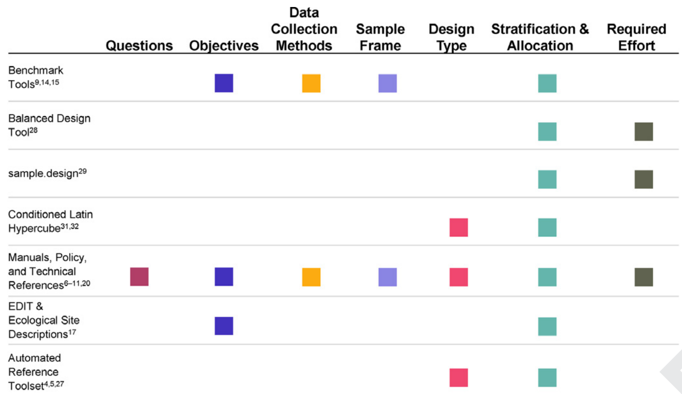

Figure 1. Example tools available and which steps in sampling design creation they apply to. Tools can inform multiple steps and each step can take advantage of multiple tools. This is not an exhaustive list of tools and—depending on the nature of a specific sampling design—some of the listed tools could be useful for unmarked steps (e.g., Ecological Site Descriptions could potentially be used to understand landscape variability as it applies to the required effort).

alize observations. Further, peer-reviewed methods are defensible in court, should the sampling effort become involved in litigation. Accordingly, modifications to standardized methods should not be made without careful consideration because they can compromise compatibility with existing data, affect accuracy or precision of indicator estimates, and expose monitoring programs to legal challenges. As an example, when evaluating the condition of sage-grouse habitat, many of the most useful methods (e.g., line-point intercept for estimating foliar cover) are described in publications specifically about sage-grouse habitat.[9]

The primary tools used to determine data collection methods are manuals, technical references, and policy sources (Fig. 1). These types of documents (e.g., MMGSSE[1], the sage-grouse HAF,[9] and *Interpreting Indicators of Rangeland Health*[20]), contain well-documented, field-tested, and peer-reviewed methods. By comparing available methods with the monitoring objectives and associated benchmarks, a small set of methods can be selected to meet the data needs. This avoids the pitfalls of creating new methods, which may not collect data with the accuracy or precision needed to be defensible in litigation. Novel methods may be appropriate, but should be considered and tested against existing methods to understand their biases and limitations.

If monitoring questions and objectives are not fully addressed with available methods, we suggest adding your new or customized protocols while keeping the core established methods intact. If selecting line-point intercept as the method for measuring foliar cover for sage-grouse habitat, then the method is described in detail in MMGSSE, which includes the full written methodology, a data recording format, and calibration protocols to ensure the data are consistent and comparable to existing data.[1]

Sample frame

A sample frame provides a sampling design with constraints, including the spatial and environmental extent of the sampling effort. Sampling frames are often spatial (i.e., study or project area) but also can be focused on a species (i.e., a population or herd) and include a temporal component. Here, we focus on spatial sampling frames, which are defined to include the locations or areas relevant to the monitoring objectives and the questions intended to be answered[6,21] (Table 1, #3). Note that in trend monitoring, repeated measures are important and the sample frame needs to define the frequency of remeasurement. The sample frame is the area (or length for linear features) within that data will be collected and to which the results of data analyses are applicable. The sample frame can be determined using the monitoring questions and objectives of the sampling effort. In other words, the sample frame must include all parts of the landscape to which the sampling effort applies and no parts of the landscape that cannot or should not be sampled. For example, if the goal is to monitor sage-grouse habitat on public lands, then the sample frame must cover at least the extent of the potential estimated habitat, but would not include nonhabitat areas or areas not to be sampled, such as privately owned land. In specific applications, like oil or gas well-pad reclamation monitoring, a sample frame for reference sampling locations may be a combination of spatial (e.g., within 2 km of the pad) and environmental criteria (e.g., having the same soil and topography as the pad) that allow objective reclamation standards to be established specific to an individual well pad.[5]

The sample frame for a sampling effort can be selected and tested with multiple tools. The reporting units defined in populating the AIMBT and policy requirements are a good starting point for defining and organizing the sample frame with all reporting units entirely falling within the sample frame. Mapping tools, like geographic information systems software, are useful for visualizing options to make decisions.

Design type

Once the monitoring objectives, data collection methods, and sample frame for a sampling effort are set, the next decision point is whether the sampling design will use probabilistic or nonprobabilistic methods of selecting sampling lo-

Please cite this article as: Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl et al., Sampling design workflows and tools to support adaptive monitoring and management, Rangelands, https://doi.org/10.1016/j.rala.2021.08.005

cations. Both approaches fill particular needs and have their applications, but each also comes with restrictions that determine which is more appropriate for given monitoring objectives.

Probabilistic sampling designs allow for results within a greater sample frame area to be inferred from the sampling locations. Probabilistic (i.e., random) designs establish a set of locations that might be sampled and assigns each site a probability of being selected before using those probabilities to pick a random subset that will be sampled. For probabilistic designs, the method of "how" to select the subset must be determined. Because these sampling locations are typically not deliberately selected for and are therefore not representative of a single, specific resource use, the data collected can be reused in multiple analyses within the sampling frame.

Most statistical methods of inference (e.g., extrapolating from individual measured sampling locations to the whole sample frame) depend on the assumption of random sampling.[21] The statistical analyses possible with probabilistic sampling offer significant advantages over nonprobabilistic sampling. Statistical analyses using monitoring data can estimate the properties of the sample frame because a random subset of locations is representative of (i.e., behaves as a stand-in for) the whole set of locations. A common application of monitoring data is to calculate the central tendency values (e.g., mean or median) of relevant indicators within the reporting units. For example, if all sage-grouse habitats in a sample frame had a chance of being sampled, the vegetation data collected from a sample of locations can also be used to estimate the percent cover of perennial grass across the entire sample frame. Further, data derived via probabilistic sampling designs can estimate variability within a sampling frame that may spur other concerns or the need for stratification to organize sampling into more uniform units. Data collected probabilistically can also be cautiously reused in multiple analyses unrelated to the original purpose of the data collection if the design fits the new monitoring objectives.[22] As an example, data collected to answer questions about rangeland health may also be used in analyses focusing on the evaluation of grazing permits.[11] However, there are constraints on reusing data, however, which depend on the details of a sampling design. In particular, the extent of the sample frame—and its overlap with other sample frames if multiple data sets are combined—sets limits on inference extent.

For a probabilistic design, the next decision is how to select sampling locations from all possible locations in the sample frame. Two common probabilistic methods for sampling landscapes are simple random and spatially balanced random (Fig. 2). A simple random design is straightforward, selecting one random location at a time from the available pool where all locations have an equal probability. One somewhat counterintuitive and often undesirable artifact of simple random sampling is the natural clustering of points.[6] Spatially balanced random sampling algorithms were developed to maintain the same statistical properties of simple random sampling but to overcome this natural clustering by spacing sample locations—for example Generalized Random Tessellation Stratified[23]—to select locations that are randomly distributed and also spatially balanced across the sample frame[24,25] (Fig. 2). There are also random sampling designs in which sample probabilities can be unequal (e.g., an accessibility weighted cost approach), but unequal probabilities make analysis and interpretation more complicated.[21]

Nonprobabilistic sampling (i.e., key area or purposeful sampling) relies on the deliberate selection of specific sampling locations, which are either targeted at a specific location of interest or assumed to be representative of the entire sample frame. However, assumptions about how widely key area data can be generalized into unsampled areas is subjective and statistical interpretations are limited to the sampled sites, but these approaches are straightforward and have been used historically by land management organizations.[7] Key areas can be an appropriate option for monitoring a specific area or permitted activity, particularly when time and funding are limited, but can be challenging to defend if conclusions are extrapolated beyond sample locations and the areas they were selected to represent. Data collected from nonprobabilistic designs cannot be extended statistically beyond the sampled locations, which limits the utility of the data collected. In some cases, retrospective demonstrations of how representative a set of key area samples are relative to a broader sample frame can be used to justify inference across a broader area. For example, exhaustive environmental datasets like digital elevation models can show that a set of samples covers a similar distribution of elevation values to those throughout a sample frame. Using these post hoc assessments depends on the assumption that the environmental dataset is closely related to variation in the collected data (e.g., elevation being related to foliar cover), but these assumptions are often challenged.

For key area sampling, there are tools to select appropriate reference locations. In addition to best professional judgment and existing data, there is the Automated Reference Toolset (ART).[5] ART uses soil and topography information in an area and selects locations that closely match to act as references.[4,5] Because the inputs are spatial rasters, there is potential for remotely sensed products like vegetation functional group mapping to be used with ART.[5,26,27] This is particularly useful in evaluating recovery from disturbances where having undisturbed, matching reference plots provides context for assessing how the recovery process is unfolding.

There are software tools available for random designs. The Balanced Design Tool (BDT)[28] is a web interface that allows the rapid drawing of spatially balanced, random sampling designs from a sample frame quickly and easily. The BDT can quickly test a sampling design by creating an interactive map of the results and offering control over features like stratification and how sampling locations are distributed (see Stratification and allocation below). The R Package sample.design has the same features as the BDT with finer control over designs, using functions from the package Spsurvey to draw spatially balanced designs.[29,30] Sample.design, however, requires the ability to write R code, and the BDT provides a graphical interface. In both cases, the designs are random and repro-

Please cite this article as: Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl et al., Sampling design workflows and tools to support adaptive monitoring and management, Rangelands, https://doi.org/10.1016/j.rala.2021.08.005

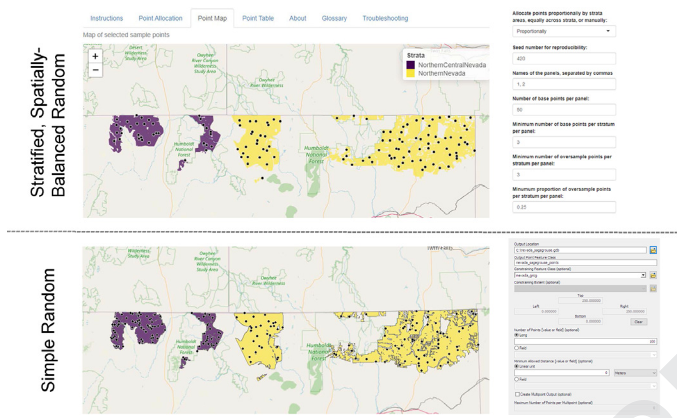

Figure 2. A comparison between two approaches to drawing a random sampling design with 100 points. The upper map shows two Sagebrush Focal Areas (SFAs) in northern Nevada with points distributed in a stratified, spatially balanced, random design and the interface from the Balanced Design Tool used to draw the points.[28] In comparison the lower map, a simple random design drawn in ArcGIS, has less even point densities across the sample frame, particularly in the Northern Central Nevada SFA (in purple). The combination of stratification and spatial balance has effectively prevented spatial clustering in the upper design, potentially increasing efficiency.

ducible because the outputs include the code and shapefiles needed to recreate the design in the future.

Stratification and allocation

Stratification divides a sample frame into smaller areas (i.e., strata) and allocates sampling effort per stratum[21] (Table 1, #7). When done correctly, stratification accomplishes three main goals for sampling design: 1) reduction in overall variance estimates by separating differing areas within the sample frame, 2) representing small yet important areas, and 3) allowing for disproportionate sampling in strata to accommodate specific monitoring objectives or to allocate more samples to strata with higher heterogeneity to better characterize those areas. Stratification is commonly used in rangeland monitoring to deal with nonuniform variability in the sample frame. If a sample design's strata are defined homogeneously (i.e., variability within strata is less than variability among strata), then the resulting increase in statistical power requires smaller sample sizes to meet monitoring objectives.

Stratification can be used to allocate samples disproportionately across the sampling frame. For example, stratification can ensure that important yet less spatially prevalent areas (e.g., riparian areas) are sampled. In a simple random or spatially balanced random design the likelihood of sampling a rare, but important area, is low because it makes up a small proportion of the sample frame. Alternatively, more sample sites could be put in strata that were expected to be sensitive to change (e.g., more human impact) and less effort put into sampling areas expected to be more stable (e.g., remote areas). Strata can be applied to a sampling design before or after determining the required effort (see Estimating required effort below), depending on monitoring objectives. Stratification should not be undertaken lightly because it introduces significant complications for analyses, including adjusting for differing variation among strata and the loss of inference area when combining data with other stratified designs.

One common goal of sampling design is to best represent the environmental variability in the sample frame, which may require stratification with multiple criteria. For example, a manager might know from experience that vegetation communities vary with elevation, which influences the characteristics of the sage-grouse habitat. Standard approaches to stratification require sampling locations be allocated to each stratum combination (e.g., high-elevation shrubland, low-elevation shrubland, and high-elevation grassland), which quickly becomes intractable.

Several recently developed sampling approaches allocate samples across heterogeneous landscapes on the basis of matching allocation patterns to the patterns in multiple environmental variables, which can approximate complex variability across a landscape.[31,32] A conditioned Latin Hypercube (cLHS) is one of those approaches used commonly in soil science, which takes into account all environmental variables of concern to condition the probability of selection for each part of the landscape while maintaining correlation structures. For example, cLHS has been used in soil mapping along with raster data of topography, expected soil type, and difficulty of access to produce sampling designs that take all three into account and selects optimal sampling locations to provide coverage while maintaining probabilistic inference.[32,33] Although the samples from a cLHS sample draw are not explicitly weighted with respect to area sampled, central tendency and variability estimates can still be calculated as long as the environmental variables represent the entire sample frame. However, samples from a cLHS draw are designed to represent the environmental variability in a sample frame and not reduce variance (as is often the case with stratified designs). Managers interested in minimizing sample variance estimates will want to include factors or variables in their analysis (e.g., soil map unit, elevation) that may explain among sample variability in the measure of interest to increase their ability to detect change or differences.

Stratification tools include the Benchmark Tools, the BDT, and sample.design.[28,29] The benchmarks in a completed

Box 1
Glossary

Accuracy: The level of confidence a measured value closely approximates the true value.
Benchmark: A set or range of values for an indicator or metric and an associated categorical classification (e.g., 15% to 25% foliar cover from sagebrush categorized as "suitable" for greater sage-grouse habitat) often tied to ecological potential.
Confidence interval: The range of plausible values for a statistic (e.g., a mean) for a selected level of confidence (e.g., 80%) in that value.
Correlation: The statistical relationship (or tendency to vary together) between two variables whether causal or not.
Key area (purposeful) sampling: Sampling in which the sample locations are selected specifically to attempt to ensure each is representative of a known portion of the sample frame.
Monitoring: The orderly collection, analysis, and interpretation of resource data to evaluate progress toward meeting management objectives. This process must be conducted over time in order to determine whether or not management objectives are being met.[39]
Monitoring objectives: Specific, quantitative statements to be evaluated by analyzing data collected as part of a sampling design.
Precision: A measure of variability of measurements around their single true value (e.g., the level of agreement between multiple measurements of a single resource).
Probabilistic (random) sampling: Sampling in which each portion of the sample frame has some chance (i.e., probability) of being selected for sampling and the sample locations are picked using a form of random selection rather than purposefully.
Reporting unit: A subset of a sample frame within which data are summarized (e.g., if the sample frame was all greater sage-grouse habitat, a reporting unit might be the breeding habitat within a specific pasture).
Sample frame: The extent of the resource (in time and space) sampling design is measuring. This is also the maximum extent to which the data from statistical analyses can be applied.
Sampling: The process of measuring a subset of a resource to estimate the properties of the resource overall (e.g., monitoring by collecting data at several specific locations to describe the landscape within which those locations fall).
Sampling design: The set of inputs, expectations, implementation decisions, and sampling rules for data collection.
Stratification: The process of breaking a sample frame into subunits (strata) to distribute the sampling between the subunits with the goal of accounting for known landscape variance (e.g., soil map units, climate zones), ensure adequate sampling in reporting units, or both.
Stratum: A spatially exclusive subset of a sample frame within which a portion of the overall sampling effort can be distributed, typically selected to reduce within-stratum heterogeneity or to ensure sampling occurs within specific areas.
Type I error: An error in which a statistical test for significance incorrectly concludes the result is significant (i.e., a "false alarm" or false change/difference error).
Type II error: An error in which a statistical test for significance incorrectly concludes the result is *not* significant (i.e., a "missed change" error).
Variance: A statistical measure of the distribution of measured values around their mean.

Benchmark Tool often take into account ecological potential or management units, both of which can suggest natural groupings of ESDs or management units to create homogeneous strata. ESD information can be found through EDIT and tied to soil maps to create strata.[17] The BDT and sample.design are both capable of handling stratified designs, the former in a point-and-click interface and the latter in R code. The ability to see the outcome of a proposed design immediately in the BDT through an interactive map makes it easier to iterate through possible stratification and allocation schemes to find one that works well (Box 1).

Estimating required effort

Once stratification and allocation decisions have been made, sampling effort (i.e., number of sample locations) needs to be allocated. Because sampling is always constrained by time, funding, and available labor, making decisions about where to apply available resources to maximize data value is important (Table 1, #8). An initial estimate of the amount of data needed must be made before any allocation occurs. Heterogeneity of the sample frame, heterogeneity within potential strata, variability of the data collection methods, and other factors have a direct influence on the minimum amount of data necessary to meet the objectives of a sampling effort. Determining what constitutes a "sufficient" sample requires significant thought and is helped by existing data, which can be used to estimate variability in the sample frame. For the example of needing to detect sagebrush cover within a 15% to 25% cover range, one would need to estimate the expected standard deviation of the cover values and decide errors can be tolerated.[34] Type I errors—so called "false alarms"—would lead to concluding sagebrush cover was in the desired range when it was not. Type II errors would lead to missing a desirable outcome or concluding cover was not within the desired range when it was. From convention, Type I error rate is typically set at 5% (i.e., $\alpha = 0.05$) and Type II error rate at 10% to 20% (corresponding to statistical power from 80% to 90%). However, the consequences of each type of error should drive the selecting of acceptable error rates.[35,36]

This approach to sample allocation brings up a paradox of sampling design in which one needs to estimate the variability of an area they need to sample before sampling it. In general, detection of smaller changes, lower Type 1 error, and higher power require more samples to achieve. In more variable landscapes, such as those with different soils and multiple intermixed ecological potentials, it may not be appropriate to assume that the variation is similar across a sampling frame. Such cases require more complicated consideration of sample sizes and use of strata to parse out variation at the outset. These situations often result in stratified designs where power analysis may be specific to the different strata used (e.g., soil types). There are also situations in which logistics, not sufficiency, are the primary limiting factor; available funding, labor, and time may set a hard limit on how much sampling can be done, but often adjusting data collection methods and using landscape stratifying variables can ensure monitoring objectives are met.[37] Although these conditions are not ideal, a realistic expectation is to conduct as much sampling as can be afforded in the available time.

Summary

The decisions required to make a sampling design can be daunting, but are manageable. They involve judgement calls, discussions with colleagues and outside experts, and repeated iteration to address the questions and understand the land-

Please cite this article as: Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl et al., Sampling design workflows and tools to support adaptive monitoring and management, Rangelands, https://doi.org/10.1016/j.rala.2021.08.005

scape at hand. There are many tools available to assist with each step of the process, several of which are described here, making the process more workable and efficient (Fig. 1). For more complicated landscapes or questions, we present a variety of new tools to stratify or optimize a sample to address the situation including the sample.design R package[29], the AIMBT[14,15], the BDT[28], cLHS[31,32], the ART[4,5], and new easily accessible soils and ecological site data[17,38] (Fig. 1). The proliferation of new spatial data and tools holds promise for producing more efficient and justifiable sampling designs. In all these cases, tools make the process both easier and better documented and the outcomes more reliable.

Declaration of competing interest

Jason W. Karl is the Editor in Chief for Rangelands but was not involved in the review or decision process for this manuscript.

References

1. Herrick JE, Van Zee JW, McCord SE, Courtright EM, Karl JW, Burkett LM. *Monitoring Manual for Grassland, Shrubland, and Savanna Ecosystems*. 2nd ed. USDA-ARS Jornada Experimental Range; 2017.
2. Nusser SM, Goebel JJ. The National Resources Inventory: a long-term multi-resource monitoring programme. *Environ Ecol Stat*. 1997; 4(3):181–204.
3. Herrick JE, Lessard VC, Spaeth KE, et al. National ecosystem assessments supported by scientific and local knowledge. *Front Ecol Environ*. 2010; 8(8):403–408.
4. Di Stéfano S, Karl JW, Bailey DW, Hale S. Evaluation of the automated reference toolset as a method to select reference plots for oil and gas reclamation on Colorado Plateau rangelands. *J Environ Manage*. 2020; 265. doi:10.1016/j.jenvman.2020.110578.
5. Nauman TW, Duniway MC, Villarreal ML, Poitras TB. Disturbance automated reference toolset (DART): assessing patterns in ecological recovery from energy development on the Colorado Plateau. *Sci Total Environ*. 2017; 584-585:476–488.
6. Elzinga CL, Salzer DW. Measuring & Monitoring Plant Populations. *US Department of the Interior*. Bureau of Land Management; 1998.
7. Coulloudon B, Eshelman K, Gianola J, et al. Utilization Studies and Residual Measurements: Interagency Technical Reference. Accessed. Q3
8. Martinson FK, Willoughby J. Statistical Considerations in Rangeland Monitoring: Technical Reference. Accessed.
9. Stiver SJ, Rinkes TE, Naugle DE, Makela PD, DA Nance, Karl JW. *Sage-Grouse Habitat Assessment Framework: Multi-Scale Habitat Assessment Tool*. Bureau of Land Management and Western Association of Fish and Wildlife Agencies; 2015 Technical Reference 6710-1.
10. Johnson DH. Monitoring that matters. In: Cooper AB, Licht DS, Millspaugh JJ, Gitzen RA, eds. *Design and Analysis of Long-Term Ecological Monitoring Studies*. Cambridge University Press; 2012:54–73.
11. Kachergis E, Lepak N, Karl S, Miller SW, Davidson Z. Guide to Using AIM and LMF Data in Land Health Evaluations and Authorizations of Permitted Uses: Technical Note 453 | Bureau of Land Management. Accessed February 10, 2021. https://www.blm.gov/documents/noc/blm-library/technical-note/guide-using-aim-and-lmf-data-land-health-evaluations-and
12. Webb NP, Kachergis E, Miller SW, et al. Indicators and benchmarks for wind erosion monitoring, assessment and management. *Ecol Indic*. 2020; 110.
13. Germino MJ, Barnard DM, Davidson BE, et al. Thresholds and hotspots for shrub restoration following a heterogeneous megafire. *Landsc Ecol*. 2018; 33(7):1177–1194.
14. Cappuccio N.*Bureau of Land Management Lotic AIM Benchmark Tool*; 2020. Accessed. https://aim.landscapetoolbox.org/analysis-reporting/analysis-and-reporting-tools/
15. Laurence-Traynor A.*Bureau of Land Management Terrestrial AIM Benchmark Tool*; 2020. Accessed. https://aim.landscapetoolbox.org/analysis-reporting/analysis-and-reporting-tools/
16. Stauffer NG, Garman SL, Karl JW. *Aim.Analysis: Analysis and Reporting of Terrestrial AIM Data Using TerrADat and Sample Design Documentation*; 2020. Accessed. https://github.com/nstauffer/aim.analysis Q4
17. Bestelmeyer BT, Williamson JC, Talbot CJ, Cates GW, Duniway MC, Brown JR. Improving the effectiveness of ecological site descriptions: general state-and-transition models and the Ecosystem Dynamics Interpretive Tool (EDIT). *Rangelands*. 2016; 38(6):329–335. doi:10.1016/j.rala.2016.10.001.
18. Karl JW, Herrick JE, Pyke DA. Monitoring protocols: options, approaches, implementation, and benefits. In: Briske DD, ed. *Rangeland Systems: Processes, Management and Challenges*. Springer Series on Environmental Management. Springer Verlag; 2017:527–567.
19. USDI Bureau of Land Management (BLM). *Terrestrial AIM Database*. Accessed January 10, 2021. http://www.landscape.blm.gov/geoportal/rest/find/document?searchText=isPartOf%3AAIM&contentType=downloadableData&start=1&max=10&f=searchpage
20. Pellant ML, Shaver PL, Pyke DA, Herrick JE. Interpreting Indicators of Rangeland Health, Version 4. Technical Reference 1734-6. US Department of the Interior, Bureau of Land Management, National Science and Technology Center; 2005.
21. Lohr SL. *Sampling: Design and Analysis*. 2nd ed. Chapman and Hall/CRC; 2009.
22. Lohr SL. Alternative survey sampling designs: sampling with multiple overlapping frames. *Surv Methodol*. 2011; 37(2):197–213.
23. Stevens L, Olsen AR. Spatially restricted surveys over time for aquatic resources. *J Agric Biol Environ Stat*. 1999; 4(4):415–428. doi:10.2307/1400499.
24. Stevens Jr DL, Olsen AR. Spatially balanced sampling of natural resources. *J Am Stat Assoc*. 2004; 99(465):262–278. doi:10.1198/016214504000000250.
25. Theobald DM, Stevens DL, White D, Urquhart NS, Olsen AR, Norman JB. Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environ Manage*. 2007;40(1):134-146. doi:10.1007/s00267-005-0199-x
26. Waller EK, Villarreal ML, Poitras TB, Nauman TW, Duniway MC. Landsat time series analysis of fractional plant cover changes on abandoned energy development sites. *Int J Appl Earth Obs Geoinformation*. 2018; 73:407–419. doi:10.1016/j.jag.2018.07.008.
27. Fick S, Nauman TW, Brungard C, Duniway MC. Evaluating natural experiments in ecology: using synthetic controls in assessments of remotely-sensed land-treatment effects. *Ecol Appl*. 2021; 31(3). doi:10.1002/eap.2264.

Please cite this article as: Nelson G. Stauffer, Michael C. Duniway, Jason W. Karl et al., Sampling design workflows and tools to support adaptive monitoring and management, Rangelands, https://doi.org/10.1016/j.rala.2021.08.005

28. Stauffer NG.*Balanced Design Tool*; 2020. Accessed. https://www.landscapetoolbox.org/balanced-design-tool/
29. Stauffer NG, Garman SL.*Sample.Design: Generate Sample Designs for Assessment, Inventory, and Monitoring Projects*; 2021. Accessed.:http://github.com/nstauffer/sample.design
30. Kincaid TM, Olsen AR, Weber MH.*Spsurvey: Spatial Survey Design and Analysis*; 2019.
31. Roudier P, Beaudette D, Hewitt A . A conditioned Latin hypercube sampling algorithm incorporating operational constraints. *Digital Soil Assessments and Beyond*; 2012:227–232.
32. Minasny B, McBratney AB. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput Geosci*. 2006; 32(9):1378–1388.
33. Ma T, Brus DJ, Zhu A-X, Zhang L, Scholten T. Comparison of conditioned Latin hypercube and feature space coverage sampling for predicting soil classes using simulation from soil maps. *Geoderma*. 2020; 370. doi:10.1016/j.geoderma.2020.114366.
34. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992; 1(3):98–101.
35. Mudge JF, Baker LF, Edge CB, Houlahan JE. Setting an optimal α that minimizes errors in null hypothesis significance tests. *PLoS One*. 2012; 7(2):e32734.
36. Mudge JF, Barrett TJ, Munkittrick KR, Houlahan JE. Negative consequences of using $\alpha = 0.05$ for environmental monitoring decisions: a case study from a decade of Canada's Environmental Effects Monitoring Program. *Environ Sci Technol*. 2012; 46(17):9249–9255.
37. Applestein C, Germino MJ, Pilliod DS, Fisk MR, Arkle RS. Appropriate sample sizes for monitoring burned pastures in sagebrush steppe: how many plots are enough, and can one size fit all? *Rangel Ecol Manag*. 2018; 71(6):721–726.
38. Ramcharan A, Hengl T, Nauman T, et al. Soil property and class maps of the conterminous United States at 100-meter spatial resolution. *Soil Sci Soc Am J*. 2018; 82(1):186–201. doi:10.2136/sssaj2017.04.0122.
39. Bedell TE. *Glossary of Terms Used in Range Management*. Fourth ed. Society for Range Management; 1998.

Authors are: USDA, Agricultural Research Service, Jornada Experimental Range, Las Cruces, NM, USA; US Geological Survey, Southwest Biological Science Center, Moab, UT, USA; Department of Forest, Rangeland, and Fire Sciences, College of Natural Resources, University of Idaho, Moscow, Idaho, USA; US Geological Survey, Southwest Biological Science Center, Moab, UT, USA. *This article was supported in part by funding from the Bureau of Land Management and the US Geological Survey Ecosystems Mission Area. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government.*