ARTICLE

Evaluation of Design-Based Sampling Options for Monitoring Stream and Wetland Extent and Distribution in California

Leila G. Lackey · Eric D. Stein

Abstract Accurate estimates of the extent and distribution of wetlands and streams are the foundation of wetland monitoring, management, restoration, and regulatory programs. Traditionally, theses estimates have relied on comprehensive mapping. However, this approach is prohibitively resource intensive over large areas, making it both impractical and statistically unreliable. Probabilistic (design-based) approaches to evaluating status and trends provide a more cost-effective alternative; however, limited information exists about the ability of various design options to meet diverse, state-level information needs such as accounting for both streams and wetlands in a single program. This study utilized simulated sampling to assess the performance of sample design options for monitoring the extent of wetlands and streams in California. Simulation results showed significantly and reliably increased precision and reduced bias with the spatially balanced, generalized random tessellation stratified (GRTS) sampling method compared to simple random sampling. In contrast, results for stratification were mixed and highly dependent on aquatic resource type and geographic region; consequently, there was no clear, broad advantage observed for stratification. This study also demonstrated the utility of a model-based approach for evaluating design options for application in other state, tribal, and regional programs.

Keywords Wetland status and trends · Wetland mapping · Optimized sampling design · Generalized random tessellation stratified sampling (GRTS) · Simple random sampling (SRS) · Stratification

L. G. Lackey · E. D. Stein (✉)
Southern California Coastal Water Research Project,
3535 Harbor Blvd., Suite 110, Costa Mesa, CA 92626, USA
e-mail: erics@sccwrp.org

L. G. Lackey
Environmental Science and Engineering Program, University of California at Los Angeles, Los Angeles, CA 90095, USA

Introduction

Wetland and stream mapping is the foundation for many regulatory, restoration and management programs, including those that support state and federal no-net loss policies (Nusser and Goebel 1997; Mitsch and Gosselink 2000) and inform decisions on compensatory mitigation (Baron et al. 2002; Clare et al. 2011). Accurate estimates of wetland and stream extent and distribution are necessary to evaluate the effectiveness of programs and policies and serve as sample frames for ambient condition surveys.

The predominant approach for evaluating the extent of streams and wetlands (here referred to jointly as aquatic resources) has been comprehensive inventory and mapping of all aquatic features; such an approach is used by the U.S. Fish and Wildlife Service (USFWS) National Wetland Inventory (NWI). Comprehensive maps are often preferred because they can provide detailed information for all locations without assumptions or inference, are easy to understand, and can readily convey information to policymakers and the public.

While comprehensive mapping is an attractive approach, it has proven inadequate for large or complex areas. Under a comprehensive approach, the entire area must be mapped in order to provide unbiased estimates of area-wide parameters, such as total wetland area or total stream length (Nusser et al. 1998; Gregoire 1999). For large geographic areas, insufficient resources frequently prevent timely completion and updating of comprehensive aquatic resource inventories (Tiner 2009; Ståhl et al. 2010). As a result, these inventories fail to provide estimates of total extent for a single point in time. For example, the NWI, begun in the 1970s by the U.S. Fish and Wildlife Service, has yet to produce a complete, national map of wetland extent (Tiner 2009). The current NWI covers less than two thirds of the country and is composed of maps produced between 1970 and the present. As a result, the NWI provides neither an estimate of current national wetland extent nor a clear mechanism for determining change with time.

In contrast to comprehensive inventories, design-based mapping uses a probabilistic approach to produce extent and trend estimates more frequently and from significantly fewer resources (Olsen and Peck 2008). We use the statistical term "design-based" to refer to probability-based sampling designs where every individual in the population has a known, non-zero probability of selection. These selection probabilities are determined by the sampling design and are, in turn, used to infer population characteristics from the selected sample. Under a design-based approach, a grid is laid out over the entire area of interest and plots are selected at random and mapped. Then, the fraction of the target area covered by the aquatic resource of interest is easily estimated by design-based inference (Gregoire 1999; Albert et al. 2010). This approach can be independent of the spatial distribution of aquatic resources and does not require a pre-existing map of aquatic resources. By mapping probabilistically, observations can be completed at a single point in time and repeated at regular intervals, enhancing ability to estimate extent and detect trends.

While probabilistic sampling and mapping cannot produce a complete map of aquatic resources, the approach can provide unbiased estimates of area-wide extent and the uncertainties in that estimate (Albert et al. 2010). For example, while the NWI has yet to map the entire country, the U.S. Fish and Wildlife Service's design-based NWI Status and Trends program (NWI-S&T) has produced five reports over the last 30 years (Dahl 2011). These reports include statistical, quantitative estimates of losses in wetland area between the 1950s and today. Similar probabilistic programs include the Minnesota Wetland Status and Trends Monitoring Program (MN-S&T), operated by the Minnesota Department of Natural Resources (MN-S&T); and the National Inventory of Landscapes in Sweden (NILS), operated by the Swedish Environmental Protection Agency (Kloiber 2010; Ståhl et al. 2010).

Evaluation of wetland extent and distribution is particularly challenging in a state as large (424,000 km^2) and diverse (13 distinct Level-III Ecoregions) as California (Omernik 2010). In addition, the California Status and Trends (S&T) program is intended to include both wetlands and streams, which have very different spatial distributions. Wetlands are often irregularly distributed based on requisite geomorphic and hydrologic settings; whereas, streams are more uniformly distributed across the landscape. Because of these challenges, a quantitative comparison of design-based sampling options is appropriate.

This study considered two major design issues for a hybrid wetland and stream S&T program; sample selection method and stratification. This work is also the first time these parameters have been rigorously evaluated for monitoring wetland and stream extent and distribution. In previous simulation work, spatially balanced sampling methodologies have

reduced sample variance compared to non-spatially balanced methods, such as simple random sampling (SRS), which can produce clustered samples (Theobald et al. 2007). Nevertheless, SRS is still commonly used, including by the NWI-S&T program, because of ease of implementation and communication of results (Dahl 2011). Systematic sampling is the simplest spatially balanced design to implement. This approach, used by the NILS program, selects sampling locations using a regularly spaced grid (Ståhl et al. 2010). However, systematic designs may align with spatial patterns in the population and unbiased variance estimation requires knowledge of the spatial variability of the population (Flores et al. 2003). Generalized random tessellation stratified (GRTS) sampling combines the advantages of SRS and systematic sampling and is used by the MN-S&T program (Kloiber 2010). GRTS provides better spatial balance than SRS by basing sample selection on a hierarchical, square grid placed over the sample area. GRTS also avoids the spatial alignment problem of systematic sampling by maintaining a random distance between adjacent points (Stevens and Olsen 2004; Deegan and Aunan 2006).

Closely related to selection method is stratification, which can be utilized to improve the accuracy and precision of sample estimates for heterogeneous areas (Jongman et al. 2006). Conceptually, stratification improves the accuracy and precision of sample estimates by dividing the population into homogeneous subsets, in effect minimizing within-stratum variability and increasing between-stratum variability. The expectation is that the homogeneous units will be better described if sampled and analyzed separately. These more precise and accurate stratum-level estimates can be aggregated to produce a more precise and accurate estimate of the whole population. However, stratification can also reduce flexibility in sampling execution and analysis. For instance, complex re-weighting procedures are required if sample estimates are required for subsets other than the sampling strata (Brus and Knotters 2008; Chen and Wei 2009). Other methods, such as spatially balanced sampling, may more easily and reliably increase the accuracy and precision of the overall estimate. Finally, stratification to improve overall precision relies heavily on accurate prior knowledge of the population, which is not always available (Kozak and Zielinski 2007). Therefore, stratification may not be necessary or appropriate if results are not required for certain subpopulations or if there is insufficient pre-existing knowledge of the population to support the stratum allocations.

This study used simulated sampling to provide empirical statistical support for probabilistic monitoring of aquatic resource extent. Simulations focused on California, explicitly explored differences between streams and wetlands, and considered the impact of design decisions on state-level program needs. Specifically, we evaluated the following questions: Can a sample design balance measurement of wetlands, which have a patchy distribution, with

measurement of streams, which are more evenly distributed? Can a probabilistic design adequately monitor rare wetland and stream types? Can the resulting sample be analyzed for all subpopulations and regions of interest? This study used multiple iteration modeling to simulate various design options and produce a statistically based evaluation that can form the basis for a recommendation to the State of California. Although the study focused on California, the approach and results should apply for any program attempting to evaluate both streams and wetlands across large, diverse areas.

Methods

General Approach

We utilized simulated sampling to evaluate sampling design elements because of its ability to provide empirical distributions of sample point estimates such as mean wetland and stream density. We used the empirical distributions to evaluate the statistical accuracy and precision of design options such as SRS vs. GRTS and stratified vs. unstratified sampling.

Geographic Databases

We based simulations on digital stream and wetland maps in California, available for 100 % and 78 % of the state, respectively (Fig. 1). For the purposes of this study, we assumed each geodatabase represented the "true" population of wetlands and streams in California. For streams, we used the National Hydrography Dataset (NHD) plus, produced by the U.S. Geological Survey and the U.S. Environmental Protection Agency. For wetlands, we utilized the NWI, split into two subsets for analysis because of a change in mapping methodology in the mid-1990s. A key step in NWI wetland mapping is production of a map of streamline
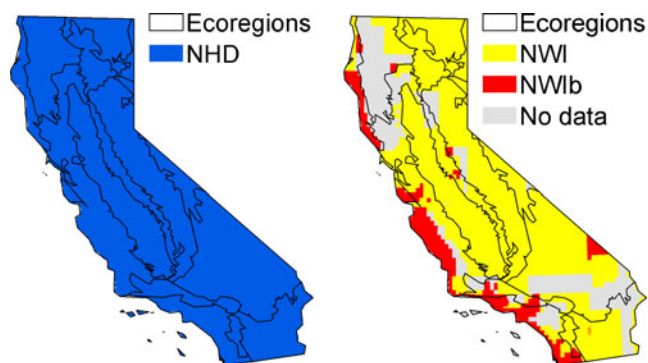

Fig. 1 Level-III ecoregion boundaries and availability of NHD and NWI digital maps in California; mapping methodology divides the NWI into maps without (NWI) and maps with (NWIb) buffered streamlines

position. Prior to the 1990s, one-dimensional features representing streamline position were kept separate from two-dimensional maps of wetland extent. However, beginning in the 1990s, one-dimensional streamlines were buffered and combined with two-dimensional wetlands into a single map of wetland and stream extent. This change in procedure significantly increased total area and altered the spatial distribution of the mapped polygons. Therefore, we considered NWI maps with buffered streamlines (NWIb), covering 10 % of California, separately from maps without buffered streamlines (NWI), covering 78 % of California.

Sampling Approaches

We considered four sampling conditions: (1) unstratified SRS; (2) stratified SRS; (3) unstratified GRTS; and (4) stratified GRTS. Numerous other sampling methods exist, but we chose to focus on SRS and GRTS and the effect of stratification. We considered spatially balanced sampling a potentially powerful mechanism for improving sample performance, as discussed earlier. However, we did not evaluate systematic sampling, another spatially balanced method and used by the NILS program, because a systematic sample of a study area cannot be easily modified for future needs. Any such modifications would require a completely new sample frame and sample draw. Other commonly employed methods, such as probability proportional to size or cluster sampling, require significant prior knowledge about the population which we could not supply (Smith et al. 2003; Kozak and Zielinski 2007). Alternative, more technical options, such as poisson sampling, were not included because they are computationally intensive without significant probability of improving sample performance (Williams et al. 2009).

We stratified along the Level-III ecoregion boundaries shown in Fig. 1 (Omernik 2010). We chose ecoregions for stratification for two primary reasons. First, ecoregions represent relatively homogenous ecological units, consistent with the assumptions of and motivations for statistical stratification. Second, aquatic resource density varied substantially between ecoregions. For example, streamline density ranges from 0.5 km km^{-2} in the Cascades to 1.2 km km^{-2} in the Southern California Mountains and wetland density ranges from 0.01 km^2 km^{-2} in the Southern California Mountains to 0.21 km^2 km^{-2} in the Northern Basin & Range. In areas covered by the NWIb maps, wetland density ranges from 0.01 km^2 km^{-2} in the Sierra Nevada to 0.59 km^2 km^{-2} in the Sonoran Basin & Range. Additionally, ecoregions are a convenient combination of numerous physical, climatological, and biological variables. These variables could be used individually for stratification, but we did not expect them to be as powerful as ecoregions. In addition, any attempt to combine variables would quickly complicate sampling and analysis. Finally, ecoregions

coincide with many of the environmental management boundaries and subunits used by the State.

When stratifying, we performed optimum allocation for variance minimization to allocate the total sample between individual strata (n_i):

$$n_i = n \left(\frac{N_i \sigma_i}{\sum_{i=1}^{k} N_i \sigma_i} \right) \quad \text{for } i \text{ in } 1, 2, \ldots, k \quad (1)$$

Under optimum allocation, total sample size (n) is allocated based on the population size (N_i) and population standard deviation (σ_i) for each stratum i. As optimum allocation is based on the presumption of a normally distributed population, we used the standard deviation of log-transformed streamline density for NHD allocation and the standard deviation of arcsine-transformed wetland density for the NWI and NWIb allocations. These transformations were selected based on the properties of the two types of variables — range from 0 to positive infinity with a right-tailed distribution for NHD and range from 0 to 1 for the NWI and NWIb.

Dataset Preparation

We began by placing a continuous, 16 km² grid over the entire state of California using the fishnet tool in ArcInfo (ESRI 2010). All grid cells were considered part of the population and the presence or extent of aquatic resources within a cell did not affect the inclusion probability. We applied a random offset (between 0 and 4,000 m) to the bottom-left corner of the grid in both the x and the y direction. We utilized the offset to reduce the probability that the fishnet tool would align grid cells with the California boundaries.

Next, we clipped the grid to the boundaries of the three geographic datasets: the state boundary for the NHD; mapped areas without buffered streamlines for the NWI; and mapped areas with buffered streamlines for the NWIb (Fig. 1). The result was a separate grid for each dataset. In addition, the area of each grid cell now represented the portion of that cell that overlapped with the mapped area. Ecoregion was defined for each grid cell based on the location of the cell centroid.

Then, we intersected the grids with NHD streamlines and NWI and NWIb polygons. Intersection split streamlines and polygons according to plot boundaries and assigned the grid cell number to each streamline and wetland segment. The numbers were then used as an index for determining the total stream length and wetland area for each grid cell for each stream and wetland subtype (defined below). Finally, we computed streamline and wetland density for each grid cell by dividing the summed lengths and areas by the cell area.

For the NHD, we considered all streamlines and five subtypes: stream order (SO) greater than 4; SO equal to 3 or 4; SO equal to 1 or 2; SO not provided; and SO equal to 1 or 2 with intermittent flow. For the NWI and NWIb, we considered all wetlands and six wetland subtypes: esutuarine, lacustrine, marine, palustrine, riverine, and palustrine, unconsolidated shore, seasonally flooded (PUSC). By including stream and wetland subtypes, we could explore sample design performance for a range of resource densities, geographic distributions, and spatial heterogeneities. In addition, these subtypes are aquatic resource groups of interest for management and research purposes in California and accurate estimate of their extent is one of the objectives of the California S&T program. The PUSC wetland subtype was used as a surrogate for rare wetland types in order to further test sampling performance (Cowardin et al. 1979). PUSC has also been used by the San Francisco Bay Area, Wetlands Regional Monitoring Program as a "classification cross-walk" to vernal pools, a unique and ecologically important wetland type in California (Holland and Jain 1981; Duffy and Kahara 2011).

Simulations

We conducted all sampling simulations in R version 2.13.1 (R Development Core Team 2011). Each of the four sampling designs was simulated 5,000 times for each dataset, a replication count used by Miller and Ambrose (2000) to give an adequate estimate of variability in the dataset. For each repetition, we recorded sample estimates of mean density of each feature type. GRTS samples were drawn using the *grts* function in the *spsurvey* package (version 2.2), developed for R and available the Comprehensive R Archive Network (CRAN) (Kincaid and Olsen 2011). SRS samples using the *sample* function in the *base* R package. We utilized random number seeds for reproducibility of GRTS and SRS sample draws.

Bias and Precision of the Sample Mean

Simulations produced empirical distributions of the mean density for each feature type and combination of sampling parameters. We utilized these empirical distributions to compare the performance of the different sampling designs. This section will describe the methods used to evaluate the empirical distribution of the mean, first to detect potential bias and second to determine the relative precision of each sampling design. Bias in the sample mean could indicate a systematic error in the sampling methodology, which oversamples a subset of the population and then fails to correct for this oversample during analysis. Improved precision (a smaller value as defined here) could indicate that the particular sample design is more reliable and a smaller sample size may be possible.

We measured bias in the sample mean by subtracting the true population value (μ) from the mean of the empirical distribution of the simulated sample means (mu_x) and dividing by the standard deviation of the empirical distribution (s_x):

$$d_{Cx} = \frac{mu_x - \mu}{s_x} \qquad (2)$$

We calculated true population values by taking the mean of all grid cells. The relationship in Eq. 2 (d_{Cx}) is known as Cohen's d and is an alternative to a t-test for the difference of means (Cohen 1988). Because our replication rate was so large (5,000), a t-test would conclude that very small differences between mu_x and μ were significant. However, Cohen's d does not consider the number of replications. Instead, the difference between the empirical distribution and the true value is only compared to the variability in the empirical distribution. Cohen's d cannot produce p-values for difference between means. However, traditional cutoffs for Cohen's d for small, medium, and large effect sizes are 0.2–0.5, 0.5–0.8, and >0.8, respectively (Cohen 1988). These cutoffs indicate that a large difference between two values is one that is close to or exceeds the variability, while a small difference is less than half of the magnitude of the variability.

We computed the precision (p_x) of each sampling design as the ratio of the standard deviation and the mean (s_x and mu_x) of the empirical distribution of the sample mean, otherwise known as the coefficient of variation:

$$p_x = \frac{s_x}{mu_x} \qquad (3)$$

We compared p_x values between sampling conditions using an f-test for the ratio of variances. This test typically has a null hypothesis that the ratio of sample variances is equal to one (i.e., $s_{x1}^2 / s_{x2}^2 = 1$). However, 1 can be replaced by any value and we chose the squared ratio of the mean of the sampling distributions (mu_x):

$$\frac{s_{x1}^2}{s_{x2}^2} = \frac{mu_{x1}^2}{mu_{x2}^2} \qquad (4)$$

Equation 4 can be re-arranged and, using Eq. 3, reduces to equality $p_{x1} / p_{x2} = 1$.

Results

SRS vs. GRTS

GRTS sample selection significantly improved precision over SRS (Fig. 2). This effect was observed for all resource types and was not affected by stratification. No method exhibited substantial bias (d_{Cx} between −0.04 and 0.02).
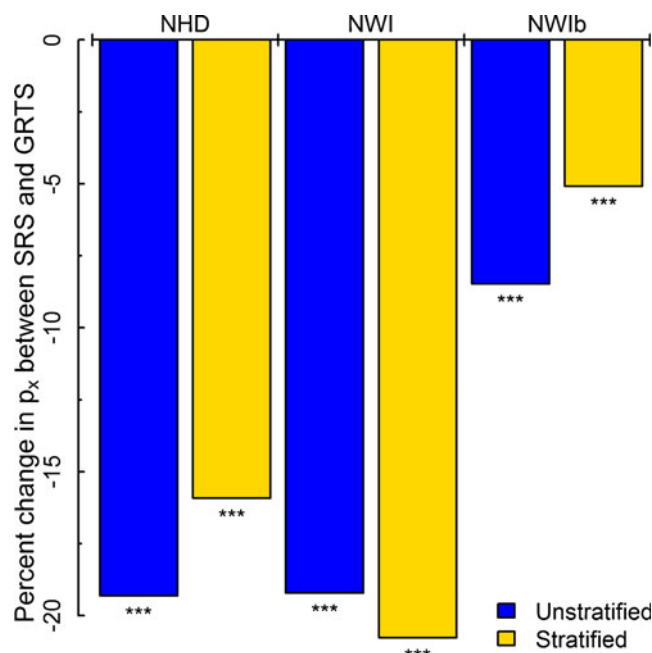

Fig. 2 Percent change in p_x between SRS and GRTS for unstratified and stratified designs for the NHD, NWI, and NWIb; *values below zero* indicate GRTS has a smaller p_x value and is more-precise than SRS; *asterisks* indicate significance level (*p-value <0.05; **p-value <0.01; ***p-value <0.001)

The observed decrease in p_x was not significantly associated with the expected spatial distribution of the resource. While the patchy NWI wetland resource had the largest percent decrease in p_x, 19 % for unstratified and 20 % for stratified designs, the evenly distributed NHD streamline resource had the second largest, 19 % for unstratified and 16 % for stratified, and the NWIb, which contains both streams and wetlands, had the smallest, 8 % for unstratified and 5 % for stratified. All differences between SRS and GRTS were statistically significant.

Stratification

The effect of stratification on sample precision was mixed (Fig. 3). Overall, stratification tended to reduce sample variance and decrease estimate sample costs; however, stratification both significantly increased and significantly decreased p_x for individual ecoregions and aquatic resources. For the NHD, stratification reduced the sample variance for all streamlines by 2.6 % under a SRS design and increased sample variance by 1.5 % under a GRTS design. For NWI and NWIb, stratification always decreased sample variance but the difference was smaller for GRTS than for SRS — 7.8 % and 18.9 % reductions for the NWI and NWIb, respectively, for an SRS design and 9.5 % and 15.9 % reductions for a GRTS design.

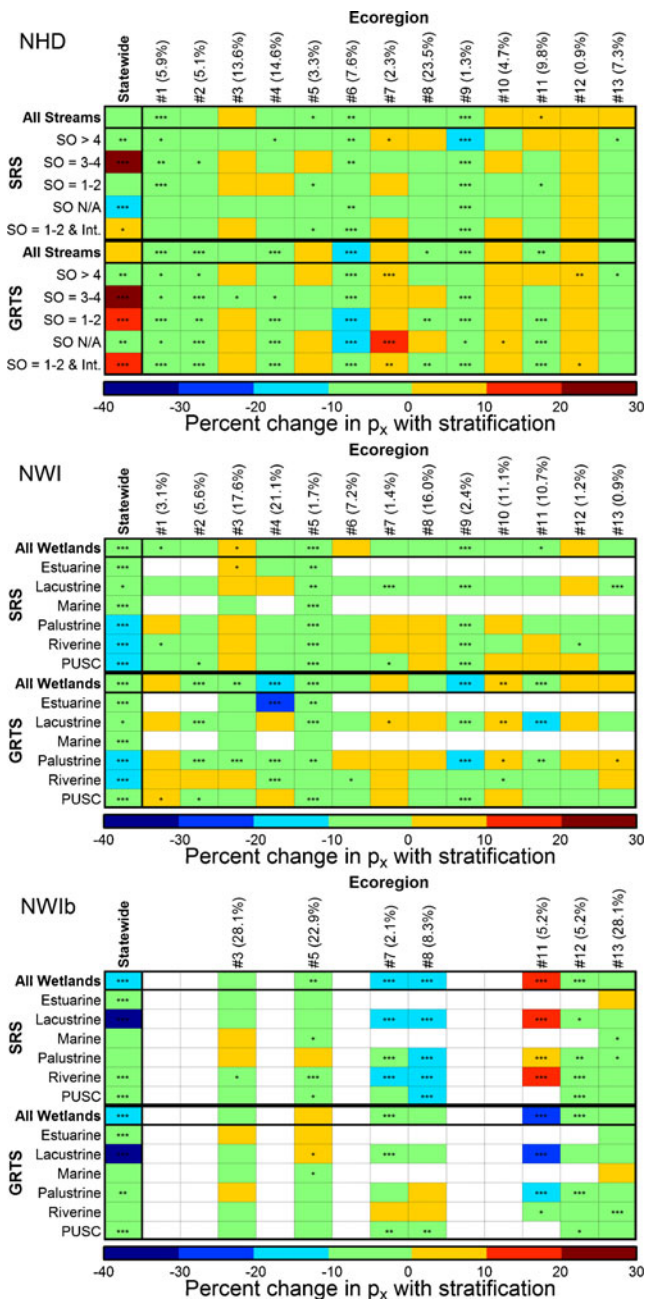The NHD and SRS sample selection illustrates the more variable effect of stratification on individual ecoregions and

Fig. 3 Percent change in p_x with stratification for SRS and GRTS designs for NHD (*top*), NWI (*middle*) and NWIb (*bottom*); *cell color indicates* the direction and magnitude of the differences; *values below zero* indicate the stratified design had a smaller p_x and was more precise than the unstratified design; *asterisks* indicate significance level (**p*-value <0.05; ***p*-value <0.01; ****p*-value <0.001); *values in parentheses* indicate the percent of the sample allocated to each of the 13 ecoregions, defined as: (*1*) Cascades; (*2*) Central Basin and Range; (*3*) Central California Foothills and Coastal Mountains; (*4*) Central California Valley; (*5*) Coast Range; (*6*) Eastern Cascades Slopes and Foothills; (*7*) Klamath Mountains/California High North Coast Range; (*8*) Mojave Basin and Range; (*9*) Northern Basin and Range; (*10*) Sierra Nevada; (*11*) Sonoran Basin and Range; (*12*) Southern California Mountains; (*13*) Southern California/Northern Baja Coast

aquatic resources. For the NHD and SRS, the effect of stratification ranged from a 11 % decrease in p_x for high-order streams in ecoregion 9 (the Northern Basin and Range) to a 3 % increase for high-order streams in ecoregion 7 (the Klamath Mountains/California High North Coast Range). For the NHD and GRTS, the range was from a 11 % decrease for streams without a recorded stream order in ecoregion 6 (the Eastern Cascades Slopes and Foothills) to a 10 % increase for streams without a recorded stream order in ecoregion 7. Similar variability was observed for the NWI and NWIb. Stratification was slightly more likely to improve precision, and decrease p_x for SRS than for GRTS, but this effect was slight. Stratification was also slightly less likely to improve precision for less common resource types, such as high-order streams in the NHD or lacustrine wetlands in the NWI, although this relationship was also highly inconsistent.

Discussion

Evaluation of design based approaches for estimating extent of streams and wetlands showed: (i) design-based sampling balanced measurement of wetlands, which have a patchy distribution, with measurement of streams, which are more evenly distributed; (ii) the GRTS approach produced less biased estimates than SRS and provided representation of all subpopulations and regions of interest; (iii) there was no clear and consistent advantage of stratification over an unstratified design; and (iv) the design-based approaches used may not adequately sample rare or spatially restricted aquatic resources.

The observed benefits from GRTS sampling are consistent with the theoretical basis of the sampling methodology and provide substantial benefits over SRS. By increasing the diversity and balance of sampled landscapes, spatially balanced sampling is expected to minimize the potential impacts of small-scale autocorrelation on sample variance (Stevens and Olsen 2003, 2004). Our results also suggest GRTS sampling may be a more effective approach, in some contexts, for reducing sample variance than use of stratification with optimum allocation. Optimum allocation reduces sample variance by allocating sample locations to individual stratum according to both the size and the variance of the population within each stratum (Bosch and Wildner 2003). This approach, used by the NWI-S&T program, can produce a spatially representative sample and reduce sample variance, but requires accurate information about the spatial variability in the population (Dahl 2011). Stratum size and variability drive stratification with optimum allocation (Bosch and Wildner 2003); therefore, stratification is most likely to reduce variance for large strata or subpopulations, and is less likely to reduce variance, or may

even increase variance, for small strata or rare subpopulations. In contrast, the size and variability of individual strata do not drive allocation for spatially balanced sampling methods, such as GRTS, and therefore may be more likely to decrease variance, as was observed here, for all strata and subpopulations.

Other simulated sampling studies have shown that stratification can be employed to reduce overall sample variance and to guarantee minimum sample sizes for subpopulations of interest (Miller and Ambrose 2000; Jongman et al. 2006). While stratification is commonly viewed as a significant improvement for many sampling approaches, we do not believe it is appropriate for the California S&T program. The mixed results for different regions and resource types do not, by themselves, provide consistent support for or against stratification. In addition, stratification may pose other limitations for implementation of an S&T program. First, ecoregions are not the only subregions the State may want to use for reporting results. Reporting may be required by geopolitical units such as counties or congressional districts. Unstratified sampling preserves the ability to conduct post-hoc analysis using a variety of groups or categories, thereby maintaining flexibility in the overall program design. Second, the allocations used in simulations were based on the stream and wetland distribution in the NWI and NHD. These allocations most likely do not represent the ideal allocation due to the incompleteness of the datasets and changes in stream and wetland extent since creation of the NHD and NWI data layers. Therefore, these allocations are unlikely to be accurate for the implemented S&T program. The simulations represent a best-case scenario where the information used for allocation is accurate and complete. If this best-case scenario cannot provide clear and consistent support for stratified over unstratified GRTS sampling, it seems less likely that an actual implementation, where the allocation information is incomplete and possibly inaccurate, will be successful.

While estimates of overall aquatic resource density are potentially achievable with acceptable levels of precision, estimates for rare or spatially limited aquatic resource types had significantly lower precision in this study. Sample variance improved for GRTS compared to SRS, but was not reliably or significantly improved by stratification. Accurate estimates for rare populations are a challenge for all probabilistic sampling designs. Options to address this issue typically lead to substantially different designs, such as adaptive sampling, stratification with regional intensification, or modification of basemaps and target regions (Smith et al. 2003; Guisan et al. 2006). Each of these designs requires assumptions about the distribution of the rare population. However, these assumptions can potentially bias the resulting estimates if based on incomplete information or if applied imperfectly (Thompson and Seber 1994). Therefore,

modification of the sampling design to address limitations in rare population measurement should only be pursued if monitoring objectives specifically emphasize accurate estimates for rare populations over other objectives. In the case of the California S&T program, the emphasis is on accurately monitoring total aquatic resource extent; therefore, modification of the design to possibly increase accuracy for a rare subtype would be inappropriate. If information about rare wetlands or streams becomes more important at a later point in time, the results from the S&T observations could possibly provide the information necessary to target rare subtypes.

Our results suggest GRTS provides both spatial balance and precision for the monitoring of aquatic resource extent and distribution. This result agrees with analysis conducted by the MN-S&T program, which also uses unstratified GRTS sampling (Kloiber 2010). The MN-S&T program also specifically evaluated the benefits of stratification according to ecoregion, but determined it did not offer statistical advantages (Deegan and Aunan 2006). In contrast, the NWI-S&T and the NILS use stratification and systematic sampling, respectively (Ståhl et al. 2010; Dahl 2011). However, both of those programs chose their respective approaches in order to increase the spatial balance and improve the precision of the resulting sample. It is also important to note that the GRTS sampling methodology was developed after the NWI-S&T program was designed.

Inclusion of both streams and wetlands is a substantial difference between the objectives of the California S&T program design and the objectives of existing S&T programs, which focus on monitoring wetland extent. In addition to hydrologic connections, California streams provide valuable aquatic habitat for a number of species that may also utilize wetland habitats (Riley et al. 2005, Collins et al. 2008). Accurate stream extent and distribution information is also relevant for many of the same scientific and management applications as wetland extent information. Our analysis shows that aquatic resources with very different spatial distributions — evidenced in California by patchy wetlands and more evenly distributed streams — can be simultaneously accommodated by a GRTS based design.

Probabilistic monitoring clearly cannot replace comprehensive maps, which are essential for site-specific actions. However, this study supports the use of probability-based monitoring of wetland and stream density as part of a coordinated strategy for monitoring wetland and stream extent and condition. In addition to providing estimates of wetland and stream density, with unbiased measures of uncertainty, probabilistic maps can serve as a sample frame for ambient, field-based condition assessment using tools such as wetland rapid assessment and indices of biotic integrity. At the time of this writing, less than ten percent of the State of California has wetland or stream maps

produced within the previous 10 years. Absence of an appropriate basemap significantly handicaps probability based investigations of wetland or stream condition using field-based methods, performed outside of recently mapped areas. Probabilistically mapped plots provide a cost-effective method for bridging this gap by providing spatially distributed, sampling units. Mapped S&T plots can provide a representative sample frame from which a subset of plots could be randomly selected and used for field-based assessments of condition or function.

Results of this study clearly show the feasibility and promise of a probabilistic approach to estimating wetland status and trends. By providing a spatially balanced sample, GRTS significantly and consistently reduced sample variance, thereby increasing power to detect change and reducing the necessary sample size. GRTS sampling also provides additional statistical and practical advantages, not directly addressed here. These advantages are a result of how the GRTS sample is drawn and analyzed. First, GRTS samples can be analyzed using a local variance estimator which reduces sample variance and, therefore, the minimum sample size (Stevens and Olsen 2003, 2004). Second, GRTS sample selection greatly simplifies the selection of an "over-sample" to provide additional locations for substitution in case locations in the original sample are unsuitable for study objectives (Larsen et al. 2008). The over-sample, also known as a master sample, approach allows for local or regional intensification and removes the need to perform a supplementary sample draw for local mapping efforts, which would require GIS and statistical software expertise as well as access to the original sample frame. Perhaps most importantly, the master sample ensures that additional sampling locations can be added over time while maintaining the spatial balance of the entire sample—as long as locations are used in order from the GRTS master sample list (Theobald et al. 2007).

Conclusions

This study supports an unstratified, GRTS sampling design for monitoring aquatic resource extent in California. Based on our simulated sampling results, this design consistently and reliably decreases sample variance compared to stratified and SRS designs. These gains are also consistent for both wetlands, which have a patchy distribution in California, and for streams, which have a more uniform distribution. Our results provide the first rigorous evaluation of applying an unstratified, GRTS design to a S&T monitoring program. These programs are a cost-effective mechanism for monitoring aquatic resource extent over large areas — necessary for scientific, management, and policy decision-making. Finally, the approach we used can easily be extended to other states and natural resources to evaluate probabilistic design questions.

References

Albert CH, Yoccoz NG, Edwards TC et al (2010) Sampling in ecology and evolution—bridging the gap between theory and practice. Ecography 33:1028–1037. doi:10.1111/j.1600-0587.2010.06421.x

Baron JS, Poff NL, Angermeier PL et al (2002) Meeting ecological and societal needs for freshwater. Ecological Applications 12:1247–1260. doi:10.1890/1051-0761(2002)012[1247:MEASNF]2.0.CO;2

Bosch V, Wildner R (2003) Optimum allocation of stratified random samples designed for multiple mean estimates and multiple observed variables. Communications in Statistics - Theory and Methods 32:1897–1909. doi:10.1081/STA-120023258

Brus DJ, Knotters M (2008) Sampling design for compliance monitoring of surface water quality: a case study in a Polder area. Water Resources Research 44:W11410. doi:10.1029/2007WR006123

Chen D, Wei H (2009) The effect of spatial autocorrelation and class proportion on the accuracy measures from different sampling designs. ISPRS Journal of Photogrammetry and Remote Sensing 64:140–150. doi:10.1016/j.isprsjprs.2008.07.004

Clare S, Krogman N, Foote L, Lemphers N (2011) Where is the avoidance in the implementation of wetland law and policy? Wetlands Ecology and Management 19:165–182. doi:10.1007/s11273-011-9209-3

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum Associates, Hillsdale

Collins JN, Stein ED, Sutula M et al (2008) California Rapid Assessment Method (CRAM) for Wetlands

Cowardin LM, Carter V, Golet FC, LaRoe ET (1979) Classification of wetlands and deepwater habitats of the United States

Dahl TE (2011) Status and trends of wetlands in the conterminous United States 2004 to 2009

Deegan G, Aunan T (2006) Survey design factors for wetland monitoring. Minnesota Department of Natural Resources, Forestry/Resource Assessment

Duffy WG, Kahara SN (2011) Wetland ecosystem services in California's Central Valley and implications for the Wetland Reserve Program. Ecological Applications 21:S18–S30. doi:10.1890/09-1338.1

ESRI (2010) ArcInfo 10. Environmental Systems Resource Institute, Redlands, California

Flores L, Martinez L, Ferrer C (2003) Systematic sample design for the estimation of spatial means. Environmetrics 14:45–61

Gregoire TG (1999) Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research 28:1429–1447

Guisan A, Broennimann O, Engler R et al (2006) Using niche-based models to improve the sampling of rare species. Conservation Biology 20:501–511. doi:10.1111/j.1523-1739.2006.00354.x

Holland R, Jain S (1981) Insular biogeography of vernal pools in the Central Valley of California. American Naturalist 117:24–37. doi:10.1086/283684

Jongman RHG, Bunce RGH, Metzger MJ et al (2006) Objectives and applications of a statistical environmental stratification of Europe. Landscape Ecology 21:409–419. doi:10.1007/s10980-005-6428-0

Kincaid TM, Olsen AR (2011) spsurvey: Spatial Survey Design and Analysis

Kloiber SM (2010) Status and trends of wetlands in Minnesota: Wetland quantity baseline

Kozak M, Zielinski A (2007) Comparison of efficiency of stratified and unequal probability sampling. Communications in Statistics - Theory and Methods 36:807–816. doi:10.1080/03610910701419695

Larsen DP, Olsen AR, Stevens DL (2008) Using a master sample to integrate stream monitoring programs. Journal of Agricultural, Biological, and Environmental Statistics 13:243–254. doi:10.1198/108571108X336593

Miller AW, Ambrose RF (2000) Sampling patchy distributions: comparison of sampling designs in rocky intertidal habitats. Marine Ecology Progress Series 196:1–14. doi:10.3354/meps196001

Mitsch WJ, Gosselink JG (2000) The value of wetlands: importance of scale and landscape setting. Ecological Economics 35:25–33

Nusser S, Goebel J (1997) The National Resources Inventory: a long-term multi-resource monitoring programme. Environmental and Ecological Statistics 4:181–204

Nusser SM, Breidt FJ, Fuller WA (1998) Design and estimation for investigating the dynamics of natural resources. Ecological Applications 8:234–245. doi:10.1890/1051-0761(1998)008[0234:DAEFIT]2.0.CO;2

Olsen AR, Peck DV (2008) Survey design and extent estimates for the Wadeable Streams Assessment. Journal of the North American Benthological Society 27:822–836. doi:10.1899/08-050.1

Omernik JM (2010) Level III ecoregions of the Conterminous United States

R Development Core Team (2011) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Riley SPD, Busteed GT, Kats LB et al (2005) Effects of urbanization on the distribution and abundance of amphibians and invasive species in Southern California streams. Conservation Biology 19:1894–1907. doi:10.1111/j.1523-1739.2005.00295.x

Smith DR, Villella RF, Lemarié DP (2003) Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environmental and Ecological Statistics 10:7–15. doi:10.1023/A:1021956617984

Ståhl G, Allard A, Esseen P-A et al (2010) National Inventory of Landscapes in Sweden (NILS)—scope, design, and experiences from establishing a multiscale biodiversity monitoring system. Environmental Monitoring and Assessment 173:579–595. doi:10.1007/s10661-010-1406-7

Stevens DL, Olsen AR (2003) Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593–610. doi:10.1002/env.606

Stevens DL, Olsen AR (2004) Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278. doi:10.1198/016214504000000250

Theobald DM, Stevens DL, White D et al (2007) Using GIS to generate spatially balanced random survey designs for natural resource applications. Environmental Management 40:134–146. doi:10.1007/s00267-005-0199-x

Thompson SK, Seber GAF (1994) Detectability in conventional and adaptive sampling. Biometrics 50:712. doi:10.2307/2532785

Tiner R (2009) Status report for the national Wetlands Inventory Program: 2009

Williams MS, Ebel ED, Wells SJ (2009) Poisson sampling: a sampling strategy for concurrently establishing freedom from disease and estimating population characteristics. Preventive Veterinary Medicine 89:34–42. doi:10.1016/j.prevetmed.2009.01.005