AGRI MAGAZINE
(International E-Magazine for Agricultural Articles)
Volume: 03, Issue: 08 (August, 2026)
Available online at http://www.agrimagazine.in
©Agri Magazine, ISSN: 3048-8656

Application of Random and Stratified Sampling in Aquatic Environmental Studies

*Mahalakshmi S[1] and Surulivel L[2]
[1]PG Scholar, Department of Aquatic Environment Management, Dr. M.G.R. Fisheries College and Research Institute, Tamil Nadu Dr. J. Jayalalithaa Fisheries University, Ponneri, Tamil Nadu, India
[2]Assistant Professor, DFEES, Dr. M.G.R. Fisheries College and Research Institute, Tamil Nadu Dr. J. Jayalalithaa Fisheries University, Ponneri, Tamil Nadu, India
*Corresponding Author's email: mahalakshmisivanantham6@gmail.com

Environmental monitoring of rivers, lakes, estuaries, and groundwater systems requires sampling designs that balance statistical rigor against the practical constraints of cost, time, and spatial heterogeneity. This paper reviews the theory and application of two foundational probability-sampling designs, simple random sampling (SRS) and stratified random sampling (StRS), in the context of aquatic environmental studies. It presents the statistical basis of each method, compares their variance and precision properties, and synthesizes evidence from water-distribution-system monitoring, lake and estuarine water-quality surveys, groundwater-level networks, and event-based river monitoring. Across the reviewed studies, stratification by hydrologically meaningful variables such as depth, distance from source, land use, and pipe characteristics consistently reduced estimation variance relative to unstratified random designs, at the cost of greater upfront design effort. The paper concludes with practical guidance for selecting and combining sampling strategies in aquatic monitoring programs.

Keywords: Sampling design; Aquatic monitoring; Random sampling; Stratified sampling; Water quality.

Introduction

Aquatic ecosystems, including rivers, lakes, reservoirs, estuaries, groundwater aquifers, and engineered water-distribution systems, are inherently heterogeneous in space and time. Physicochemical and biological parameters such as dissolved oxygen, turbidity, nutrient concentration, and contaminant load vary with depth, proximity to pollution sources, flow regime, and season. Because it is rarely feasible to measure an entire water body exhaustively, environmental scientists rely on probability sampling to obtain statistically defensible estimates of water-quality parameters from a manageable subset of locations and times.

Sampling design directly determines the reliability, cost-efficiency, and regulatory defensibility of an aquatic monitoring program. Two designs dominate the literature and practice: simple random sampling, in which every sampling unit has an equal and independent chance of selection, and stratified random sampling, in which the target population is first partitioned into relatively homogeneous subgroups (strata) before random selection occurs within each subgroup. This paper examines both designs as applied to aquatic environmental studies, drawing on published monitoring-network research to illustrate their comparative performance.

Statistical Foundations

Simple Random Sampling

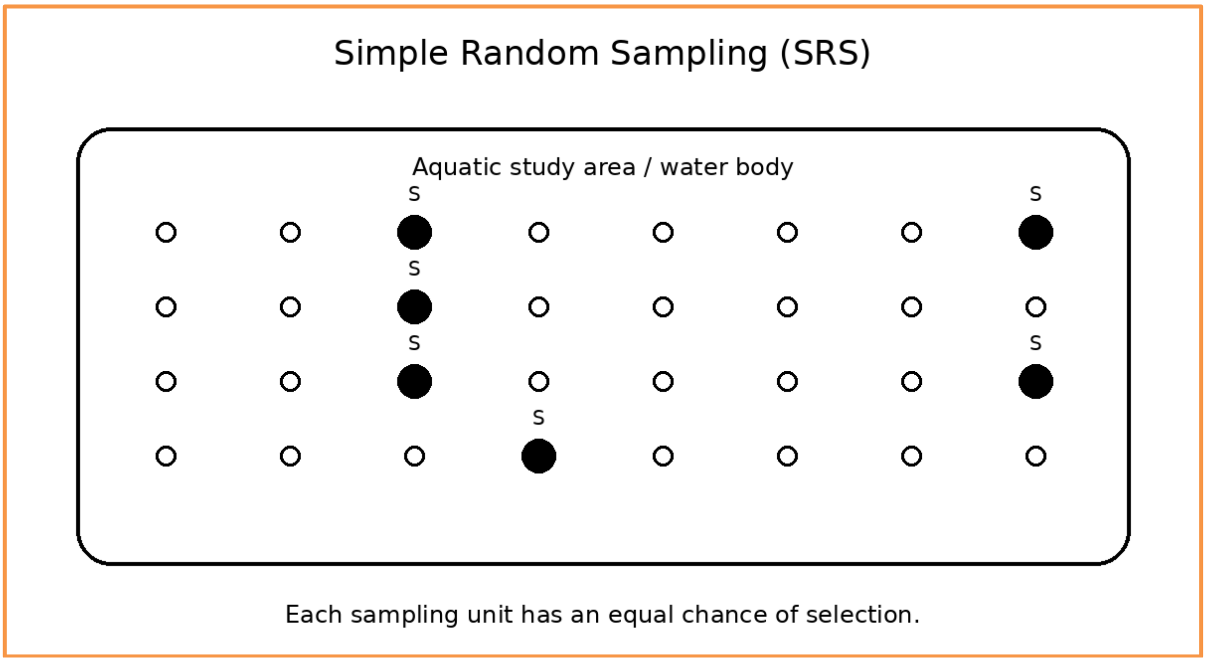

Figure 1. Simple random sampling of an aquatic study area

In SRS, n sampling units are drawn from a population of N units such that every possible combination of n units has an equal probability of selection. For a water-quality parameter y, the sample mean $\bar{y}$ is an unbiased estimator of the population mean, and its variance is $Var(\bar{y}) = S^2/n \times (1 - n/N)$, where S^2 is the population variance. SRS is easy to implement and requires no prior stratification information, but its precision is entirely governed by the total variance of the target parameter across the whole water body. When a lake, river, or aquifer is spatially heterogeneous, this total variance can be large, and SRS may under-sample rare but important zones such as pollution hot spots, tributary mouths, or shallow littoral habitats.

Stratified Random Sampling

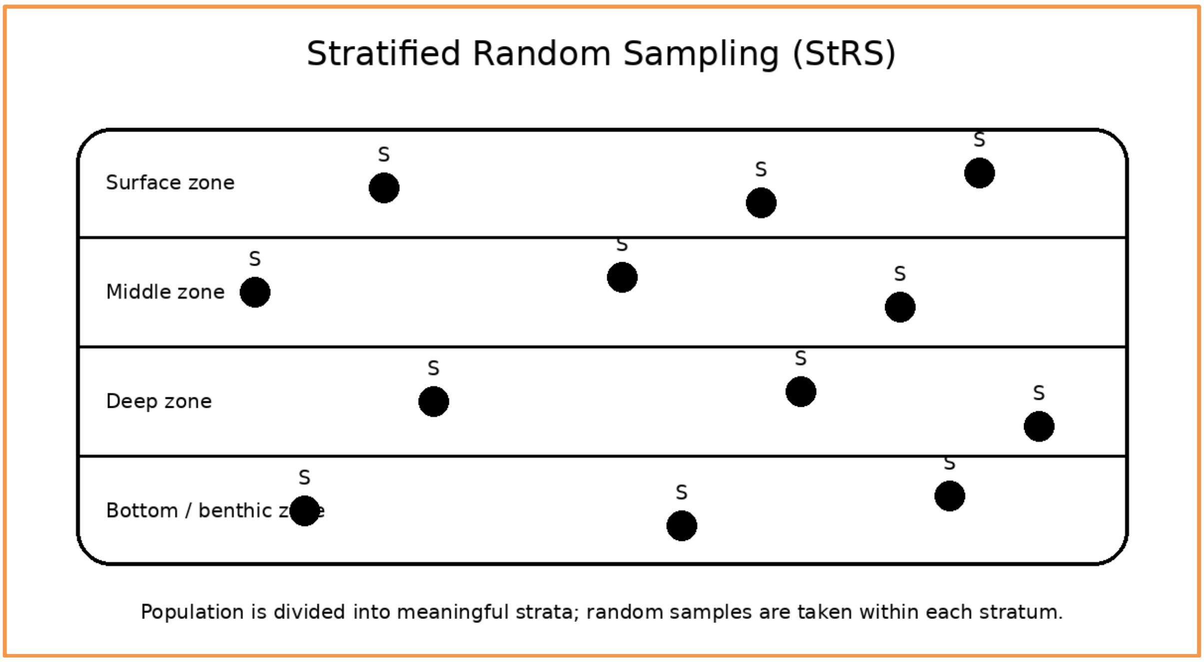

Figure 2. Stratified random sampling across depth-based aquatic zones.

StRS partitions the population into L strata that are internally homogeneous with respect to the variable of interest, then draws an independent random sample from each stratum. The stratified estimator of the population mean is a weighted average of stratum means, $\bar{y}st = \Sigma (Nh/N)\, \bar{y}h$, and its variance, $Var(\bar{y}st) = \Sigma (Nh/N)^2 (Sh^2/nh)(1 - nh/Nh)$, depends only on within-stratum variance Sh^2, not on the total population variance. Because between-stratum variability is effectively removed from the sampling error, StRS is more precise than SRS of equal sample size whenever the chosen strata capture real differences in the parameter being measured. Allocation of the sample across strata can be proportional to stratum size, or optimized (Neyman allocation) to minimize variance for a fixed budget by allocating more samples to strata with higher internal variability or higher sampling cost.

Table 1. Comparison of simple random and stratified random sampling for aquatic monitoring.

Criterion	Simple Random Sampling	Stratified Sampling
Basis of selection	Every sampling unit in the water body has an equal probability of selection	Water body divided into homogeneous strata (depth, substrate, salinity zone); units drawn independently within each stratum
Precision for heterogeneous systems	Lower; variance driven by whole-system heterogeneity	Higher; between-stratum variance is removed from the sampling error
Field logistics	Simple to implement; minimal prior knowledge required	Requires prior delineation of strata (bathymetric, hydrological, or land-use data)
Risk of spatial/temporal bias	Higher; rare habitats or events can be missed by chance	Lower; every stratum is guaranteed representation
Typical aquatic application	Preliminary reconnaissance surveys, baseline contaminant screening	Long-term monitoring networks, regulatory compliance sampling, lake/river zonation studies

Stratification Variables in Aquatic Systems

The effectiveness of StRS depends on choosing stratification variables that are correlated with the parameter of interest and can be determined prior to sampling, typically from bathymetric maps, hydrological data, land-use records, or infrastructure attributes. Commonly used strata in aquatic environmental studies include:

- Depth zones - surface, thermocline, and hypolimnion layers in lakes and reservoirs, which differ markedly in temperature, oxygen, and nutrient profiles.
- Distance from pollution source or treatment plant - used in river and water-distribution-system studies, where water quality changes systematically with travel distance and residence time.
- Substrate and habitat type - littoral versus pelagic zones, sediment type, and vegetation cover in lake and estuarine ecosystems.
- Hydrological regime or land use in the contributing catchment - agricultural, urban, or forested subbasins that generate different pollutant loadings.
- Infrastructure attributes - pipe material, pipe diameter, and age in engineered water-distribution networks, which influence corrosion, disinfectant decay, and contaminant re-growth.
- Spatial clusters derived from statistical methods - for example, k-means clustering of groundwater observation wells based on depth-to-water behavior, used to define strata for regional groundwater monitoring networks.

Applications in Aquatic Environmental Studies

Water-Distribution-System Monitoring

Studies of drinking-water distribution networks have applied randomized stratified sampling using pipe material, pipe diameter, and distance from the treatment plant as stratification variables. Distance from the treatment plant has been identified as the single most effective stratification variable for reducing estimation error in these systems, because residual disinfectant and byproduct concentrations change systematically along the flow path. Such designs also support cost-benefit analysis across utilities of different sizes, helping small, medium, and large municipalities allocate limited monitoring budgets efficiently.

Lake and Reservoir Water-Quality Surveys

In lentic systems, stratified designs based on depth and distance from shore have revealed statistically significant differences in water-quality parameters between strata, confirming that vertical and horizontal zonation are important sources of variance that unstratified sampling would obscure. Remote-sensing-assisted approaches, which combine satellite-derived water-quality proxies with spatial simulated annealing, have further shown that among random, systematic, and stratified strategies, purely random designs tend to produce the greatest sampling uncertainty for a given number of samples unless the water body is sampled very densely.

River and Event-Based Monitoring

River monitoring programs have used stratified random sampling to allocate sampling effort across subbasins grouped into strata by hydrological characteristics, enabling both regulatory-style statistical surveys and intensive process studies of in-stream kinetic transformations within selected subbasins. Event-based approaches extend this logic temporally, using post-event stratified random sampling schemes with automatic samplers to capture storm-driven pollutant pulses that would be missed by fixed-interval or purely random time sampling.

Groundwater Monitoring Networks

For regional groundwater-level monitoring, observation wells have been grouped into statistically homogeneous clusters using unsupervised clustering algorithms, with cluster membership then serving as the stratification scheme for comparing cluster random sampling and stratified random sampling. Stratified random sampling was found to be the more robust and appropriate technique for regional-scale monitoring-network design, producing more accurate spatial predictions of depth-to-water at a given sampling intensity than cluster random sampling.

Methodological Considerations and Limitations

Selecting between random and stratified designs, or combining them, requires attention to several practical factors. First, stratification is only beneficial when strata are genuinely more homogeneous internally than the population as a whole; poorly chosen strata add design complexity without improving precision. Second, prior data, whether historical monitoring records, bathymetric surveys, remote-sensing imagery, or hydrological models, are needed to delineate strata before sampling begins, which raises the upfront cost of a stratified program relative to SRS. Third, allocation strategy matters: proportional allocation is simple and ensures every stratum is represented in proportion to its size, whereas optimal (Neyman) allocation can further reduce variance but requires reliable estimates of within-stratum variance and cost, which are not always available at the design stage. Finally, both designs assume that sampling units within the frame are correctly enumerated; incomplete or outdated hydrological or infrastructure data can introduce coverage bias regardless of the sampling design chosen.

Hybrid and adaptive designs are increasingly used in practice. Two-stage designs combine stratification at a coarse spatial scale (for example, subbasins or depth zones) with random or systematic sampling within each stratum. Remote-sensing-guided designs use satellite or aerial imagery to pre-screen spatial variability and refine stratum boundaries before field sampling, reducing the number of in-situ samples needed to achieve a target precision. Clustering-based stratification, using algorithms such as k-means on historical monitoring data, offers a data-driven alternative to expert-defined strata, particularly for large or poorly characterized systems.

Conclusion

Random and stratified sampling designs each have a defined role in aquatic environmental monitoring. Simple random sampling remains a valid and low-overhead choice for reconnaissance surveys or relatively homogeneous water bodies, while stratified random sampling consistently improves the precision of water-quality and hydrological estimates in systems with well-defined spatial or infrastructural heterogeneity, including drinking-water distribution networks, lakes, rivers, and groundwater aquifers. The evidence reviewed here indicates that the choice of stratification variable, whether depth, distance from source, land use, pipe characteristics, or statistically derived clusters, is often more consequential than the choice between random and stratified sampling in the abstract. Effective aquatic monitoring programs therefore benefit from investing in the hydrological, geospatial, or infrastructural knowledge needed to define meaningful strata, and from combining stratified spatial designs with random or event-based temporal sampling to capture both persistent gradients and episodic pollution events.

References

1. Brus, D. J., & Knotters, M. (2008). Sampling design for compliance monitoring of surface water quality: A case study in a polder area. Water Resources Research, 44(11).
2. Cochran, W. G. (1977). Sampling techniques (3rd ed.). John Wiley & Sons.
3. Environmental Protection Agency (EPA). (2000). National primary drinking water regulations. Code of Federal Regulations, Title 40, Chapter 1, Part 141. U.S. Government Printing Office.
4. Harmancioglu, N. B., Fistikoglu, O., Ozkul, S. D., Singh, V. P., & Alpaslan, M. N. (1999). Water quality monitoring network design. Kluwer Academic Publishers.
5. Kish, L. (1995). Survey sampling. Wiley.
6. Rossman, L. A. (2000). EPANET 2 users manual. National Risk Management Research Laboratory, U.S. Environmental Protection Agency.
7. Sanders, T. G., Ward, R. C., Loftis, J. C., Steele, T. D., Adrien, D. D., & Yevjevich, V. (1983). Design of networks for monitoring water quality. Water Resources Publications.
8. Speight, V. L., & Kalsbeek, W. D. (2004). Randomized stratified sampling methodology for water quality in distribution systems. Journal of Water Resources Planning and Management, 130(4), 330–338.
9. Thomas, R. B., & Lewis, J. (1995). An evaluation of flow-stratified sampling for estimating suspended sediment loads. Journal of Hydrology, 170(1–4), 27–45.
10. Ward, R. C., Loftis, J. C., & McBride, G. B. (1990). Design of water quality monitoring systems. Van Nostrand Reinhold.
11. Zhang, Y., et al. (2021). Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated approach. Science of the Total Environment, 777, 146131.
12. Groundwater monitoring network design using stratified and cluster random sampling of depth-to-water clusters. (2019). Journal of Hydrology, 572, 1183–1200.
13. Event-based water quality monitoring using stratified random sampling designs. (2012). [Conference/technical report].
14. Numbers Analytics. (2025). Mastering stratified sampling: Applications in environmental science. Retrieved August 27, 2026, from https://www.numberanalytics.com/blog/ultimate-guide-stratified-sampling-field-methods-environmental-science