Forest Service
U.S. DEPARTMENT OF AGRICULTURE

Pacific Northwest Research Station | Research Note PNW-RN-584 | August 2026

Algorithmic Optimization of Sampling for the National Visitor Use Monitoring Program

Noah Creany, Eric M. White, Sarah Cline[1]

Abstract

Creany, Noah; White, Eric M.; Cline, Sarah. 2026. Algorithmic optimization of sampling for the National Visitor Use Monitoring program. Res. Note. PNW-RN-584. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 21 p. https://doi.org/10.2737/pnw-rn-584.

The National Visitor Use Monitoring (NVUM) program provides the U.S. Department of Agriculture, Forest Service and the U.S. Congress with statistically sound estimates of recreation use in the National Forest System and the economic contribution of those recreation visits to rural communities. After 25 years and five rounds of NVUM sampling, the program has assembled long-term monitoring datasets to characterize forest visitors and understand spatial and temporal patterns of use. However, because the visitor-intercept sampling methodology for NVUM surveys is cost-intensive, program managers are exploring approaches to reduce data-collection costs while maintaining the quality and utility of NVUM data. Here we describe an algorithmic optimization approach that uses historical NVUM data to generate sampling scenarios that systematically reduce the number of site-days (sampling days) while minimizing the impact on the data quality and its applications and end uses.

KEYWORDS:
Recreation
Sampling methodology
Sample optimization
Long-term monitoring
Visitor surveys

[1] **Noah Creany** is an Oak Ridge Institute for Science and Education postdoctoral research fellow, and **Eric M. White** is a research social scientist, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3200 SW Jefferson Way, Corvallis, OR 97331; **Sarah Cline** is an economist for the National Visitor Use Monitoring program, 1400 Independence Avenue SW, Washington, DC 20250-1125.

Introduction

The National Visitor Use Monitoring (NVUM) program is the U.S. Department of Agriculture, Forest Service's primary source of data and information to monitor the volume and quality of recreation use of National Forest System lands. This fulfills directives from Executive Order 12862,[2] the Government Performance and Results Act of 1993,[3] the Forest Service's 2012 Planning Rule,[4] and the EXPLORE Act.[5] The Forest Service developed and piloted NVUM in 1996 (Zarnoch et al. 2002), and implemented the program in 2000 (English et al. 2002) to provide a systematic and statistically valid methodology to estimate national forest recreation visitation to report to Congress. The NVUM program is notable among federal land management agencies because of its long-term and consistent methodology for a nationwide visitor use monitoring dataset.

In addition to estimating national forest visitation, the NVUM program conducts survey interviews from a stratified sample of national forest visitors to produce descriptive information about visitors and their visits, spending, and evaluations of the level of satisfaction from their national forest visits. Forest Service NVUM sampling is organized by rounds; in a given year, NVUM collects data at about 20 percent of the 154 national forests and grasslands. NVUM data are then incorporated into national forest monitoring and planning processes and provide a basis for evaluating the performance of the Forest Service in carrying out its multiple-use mandate and agency mission.

However, operationalizing NVUM data collection under the current sampling methodology is cost intensive. In response to tightening operating budgets within the Forest Service, the agency has directed research to investigate methods and approaches that would reduce NVUM data-collection costs, while maintaining the quality and fidelity of the data collected. Here we describe a sampling optimization methodology to systematically reduce NVUM data-collection costs and a machine learning classification model trained on the optimization results to generalize optimization results to future rounds of data collection.

Previous NVUM Efficiency Efforts and Research

The NVUM sampling methodology has remained relatively consistent since its initial round of data collection. However, program managers have introduced modifications and directed research aimed at reducing the costs and burden of data collection. For example, starting in round five of NVUM data collection (2020–2024), the Forest Service eliminated the use of automated pneumatic-tube counters and trail counters for measuring vehicle and pedestrian use, respectively. Instead, Forest Service researchers, leveraging the extensive amount of data collected in previous rounds, developed a beta-distribution approach to estimate the 24-hour visitor use at a site that required only keeping a hand tally of exiting people and vehicles during a sampling period (English et al. 2018, Zarnoch et al. 2005). This modification not only provided cost savings from eliminating the need to purchase and maintain the equipment but also led to substantial reductions in labor costs for technicians to set up and relocate counters throughout the sampling period.

[2] **Presidential Documents. 1993.** Executive Order 12862. Setting customer service standards. Federal Register. 58: 176.

[3] **Government Performance and Results Act of 1993**; 31 U.S.C. 1101 et seq.

[4] **National Forest System Land Management Planning.** 2012. 77 FR 21162. Federal Register. 77: 68.

[5] **Expanding Public Lands Outdoor Recreation Experiences (EXPLORE) Act. 2025.** Public Law 118-234. 138 Stat. 2836.

During round three of NVUM data collection (2010–2014), Askew et al. (2014) explored the challenge of reducing NVUM sampling intensity through simulations of sets of sample reduction scenarios based on NVUM data from 2011. The objective of the study was to identify a manageable set of reduction strategies that could be systematically assigned to different national forests, rather than developing custom solutions for each forest. These simulations evaluated strategies by balancing the percentage reduction in sampling days (hereafter, site-days) (with a target of an average reduction of 15–25 percent per forest, about 20 percent overall) with the impact on the reliability of visitation estimates (i.e., site visits and national forest visits mean, standard error, and coefficient of variation). This simulation-based approach aimed to provide a practical and computable method for managing budget **constraints** without making arbitrary cuts that would compromise the validity or quality of the data for long-term trend analyses.

Optimization of Sampling

The fundamental goal of a statistical sample is to collect data from a subset that is representative of the broader population of interest (Lohr 2021, Neyman 1934). However, the design of any sampling scheme requires balancing statistical rigor with practical constraints, such as budgetary limitations and logistical challenges like the accessibility of sample locations (Tillé and Wilhelm 2017). When a sampling problem involves multiple and conflicting objectives, such as minimizing costs, while maximizing the quality and representativeness of the data, a multi-objective optimization (MOO) analysis is well suited to identify a set of solutions that balance the desired objectives and constraints (Rashed et al. 2024).

A strength of MOO is that the result is not a single solution, but rather it identifies a set of Pareto-optimal solutions. Each of these solutions represents an optimal tradeoff, meaning no single objective can be improved without sacrificing performance in another (Nayak 2020). This can be advantageous to decision makers because they can then select the solution that is best suited to available resources and aligned with their priorities. While the use of MOO in the field of natural resources is less commonplace than it is in engineering and business logistics, MOO approaches have been effectively applied in natural resource fields to systematically enhance sampling efficiency. For example, Li et al. (2022) used MOO based on Pareto optimality to develop optimized sampling designs for soil mapping. Similarly, MOO has been used to generate operational sampling plans for areas that are difficult to access, a common challenge in environmental monitoring (Dumont et al. 2024). Solving these complex optimization problems often relies on the efficiency of nature-inspired computational techniques like genetic algorithms (Yang 2021), which are adept at exploring vast objective search spaces to identify the optimal set of solutions.

Methods

MOO offers a robust and algorithmic, data-driven approach to addressing challenges that the NVUM program faces because it is highly effective and efficient at identifying solutions that satisfy the objectives in the analysis. Nevertheless, the primary challenge in applying MOO to the NVUM sampling is operationalizing objectives that can be difficult to define or summarize in a single metric (e.g., representativeness of a survey sample), yet this proxy must be robust enough to ensure that the reduced sample maintains the quality of the full sample. The NVUM Scientific Advisory Committee, a panel of Forest Service subject matter experts, identified three key objectives to optimize in this analysis:

- Reduce costs associated with the number of site-days and travel to sites.
- Ensure consistency with recreation use estimates resulting from the full sample.
- Preserve representativeness of the sample population.

The unit of analysis for the optimization problem is the site-days, which is associated with a sample stratum combining the **site type** (i.e., overnight use developed site, day use developed site, general forest area, and wilderness) and the **use level** (i.e., low, medium, high, very high).

The first objective focuses on minimizing the logistical costs associated with travel to and from NVUM sites, which is a significant component of the overall cost of NVUM data collection. The second objective aims to ensure that the reduced sampling design yields generally unchanged visitation estimates. The third objective, which proved most challenging to distill into a single measure, seeks to preserve the representativeness of the reduced sample to the full sample of forest visitors. After considering a range of descriptive characteristics in the dataset, we selected the distribution of visitor activity types as a proxy for representativeness because it is a key descriptive characteristic collected in NVUM surveys. Visitor activity types provide a useful grouping variable to understand diverse types of forest visitors and their experiences.

For each objective, we developed a quantitative metric to use in the optimization analysis:

- **Travel cost**—the total travel distance cost associated with selected sampling sites, which were calculated based on the mean distance to the three nearest ranger district cities or towns from the sample site, multiplied by the federal travel cost per mile using the 2024 reimbursement rate of $0.67 (U.S. General Services Administration 2025), multiplied by the number of sampling at each site.
- **Visitation pattern consistency**—the median absolute deviation (MAD) of hand-tally counts of exiting traffic (completed by the on-site interviewers) on the site-day between the full dataset and the reduced sample, aggregated by site type and sampling stratum.
- **Sample representativeness**—the **Jaccard distance**, or complement of Jaccard similarity (1−similarity), between the set of main activity types reported by interviewed visitors in the full dataset and the reduced sample; this metric quantifies the dissimilarity or difference between two sets and provides a measure of the number of elements the sets have in common, relative to the total unique elements (Hancock 2014).

The aim of these metrics is to systematically reduce the number of required site-days, while simultaneously meeting the corresponding objective.

Furthermore, given that there is likely to be an indefinite number of solutions that satisfy the objectives, we included constraints in the optimization analysis:

- **Minimum site-days per stratum**—ensures that each stratum (defined by site type and use level) maintains a minimum required number of site-days (i.e., 8 site-days per stratum), preventing undersampling in critical categories.
- **Target accuracy**—maintains a high correlation (e.g., a target Pearson correlation, r, of 0.80) between the mean hand-tally counts of the full sample and the reduced sample within sampling strata.

These constraints circumscribe the range of potential solutions and increase the speed of convergence toward the global optima, while ensuring the feasibility and statistical validity of the reduced sampling designs.

This MOO framework provides a systematic and robust exploration of a broad range of sample reduction scenarios (e.g., 10 percent per stratum, 20 percent per stratum, or stratum-specific reductions in site-days) that can be applied to the specific composition of site types and use levels within a national forest. Furthermore, it is possible to develop a custom sample reduction scheme for any given national forest unit or subunit that balances the desired sample coverage and intensity with available resources.

To demonstrate the process of using MOO to identify NVUM sample reduction strategies, we applied this framework to the Superior National Forest for fiscal year 2026, round six of the NVUM sampling calendar (fig. 1). The Superior National Forest has participated in NVUM data collection since the program's inception in 2000 but, like many units, faced disruptions to data collection in round five (fiscal year 2021) because of the COVID-19 pandemic. For fiscal year 2026, Superior National Forest managers anticipated challenges with data collection because of limited resources and consulted NVUM program managers to identify strategies to reduce the total number of site-days for the sampling calendar. Because of the disruptions in round five (fiscal year 2021), we used NVUM data from round four (fiscal year 2016) for the optimization to identify sampling reduction strategies. However, as part of NVUM sampling methodology, before the development of the sampling calendar, national forest staff participate in the inventory design process to update the NVUM sampling sites to be sampled in the forthcoming round and substitute or change sites as their level of development changes or sites are closed. Consequently, the intersection of NVUM sites common to both fiscal years 2016 and 2026 was limited, with about 50 percent of sites common to both years. With a higher proportion of sites common to both sample years, the optimization solutions from fiscal year 2016 could be directly applied to the fiscal year 2026 calendar. However, because of this marginal overlap between sample years, the optimization solutions from fiscal year 2016 could not be directly applied to the fiscal year 2026 calendar. To address this challenge of generalizing the reduced sample scenario solutions to future rounds of data collection at unobserved sites, we developed a classification model to learn the characteristics of site-days that were identified by the optimization to drop from the sampling calendar with minimal impact on data quality. This classification model can then be used to project optimization solutions for future rounds of data collection.

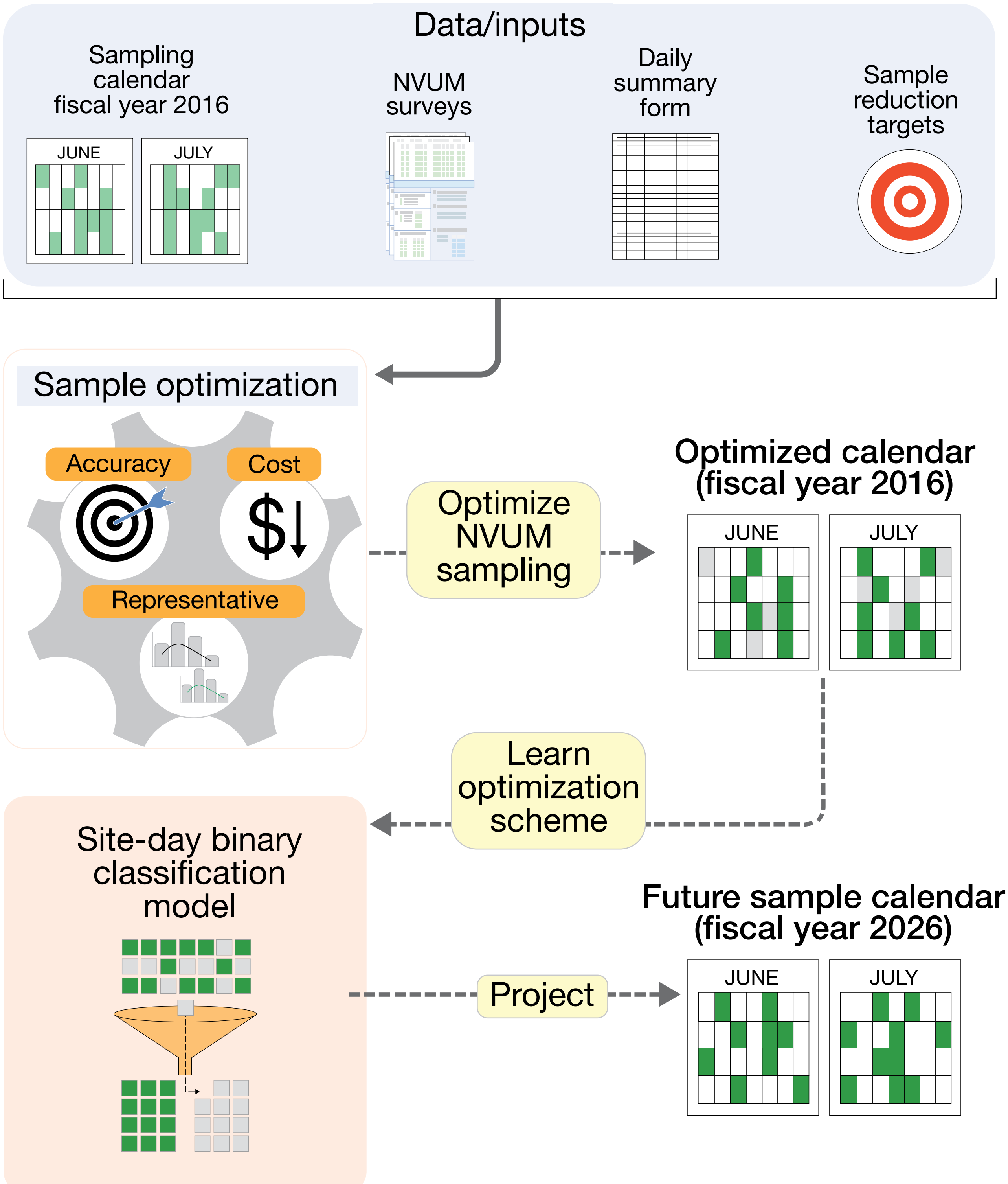

Figure 1—Analysis workflow used to identify National Visitor Use Monitoring (NVUM) sample reduction strategies at Superior National Forest for fiscal year 2026, including the data inputs to the multi-objective optimization, classification model, and sample calendar outputs. One site-day is equivalent to one sample day.

Data Preparation

Multi-Objective Optimization Framework

The primary data sources for this analysis were the NVUM dataset, consisting of individual visitor survey responses, and the **daily summary form**, which summarizes the sampling activities, such as the number of surveys collected and total hand-tally counts of **last-exiting** recreationists/vehicles. We supplemented these datasets with geospatial data files with site metadata to extract key attributes, such as site type, ranger district, and whether the site was a **proxy site** or **nonproxy site** for NVUM data collection. We excluded the proxy and **view corridor sites** from this analysis to focus on the nonproxy days that represent most site-days and require technicians to interview visitors for data collection. We standardized all datasets to ensure consistency of site names and site identifiers between datasets. Finally, we used the daily summary form, where NVUM technicians record the number of surveys collected and total hand-tally counts of last-exiting recreationists/vehicles, to construct a definitive sample frame because it provided the most accurate record of data collection.

Each unique site-day in this frame (i.e., site-id and sampling date) was assigned a unique hexadecimal identifier (hexId) to serve as a discrete decision variable to be included or excluded in the optimization model. Additionally, we constructed two subsets of the survey response dataset for use in the optimization analysis. The first was an activity-type dataset that contained the primary activity type of the survey respondent and the associated hexadecimal identifier. This dataset would be used to calculate the Jaccard distance loss between the full sample and reduced sample scenarios. In the second dataset, we summarized the data by calculating the mean end-hand-tally count for each sample stratum before determining the loss in MAD between the full and reduced sample scenarios. Finally, we summarized the total number of site-days for each stratum in the full sample frame to serve as the starting point for reduction scenarios as well as the minimum number of site-days per stratum for reduction scenarios.

Travel Cost Calculation

To incorporate the logistical costs of data collection, we modeled the travel cost to sample sites using the OSMnx (Boeing 2025) Python library. First, we queried the network of drivable roads from the OpenStreetMap[6] online application (OpenStreetMap Foundation, n.d.) within the bounding box of the Superior National Forest, which we expanded by ±3 degrees to cover the full extent of the Superior National Forest and surrounding areas (fig. 2). To increase the **accuracy** of network routing and travel time estimates, we enhanced the road graph network with attributes such as grade, elevation, and posted speed limits.

We then queried OpenStreetMap for the place nodes of the cities and towns in which ranger district offices were located for the Superior National Forest: Aurora, Cook, Duluth, Ely, Grand Marais, and Tofte. For each NVUM sampling site, we calculated the shortest path, weighted by travel time, to each ranger district office. Next, we calculated the mean travel distance from each NVUM sampling site to the nearest three ranger district offices, which served as a proxy for the logistical effort required to sample each site. We converted this travel distance into a cost by multiplying the travel distance by the federal reimbursement rate.

[6] The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Department of Agriculture of any product or service.

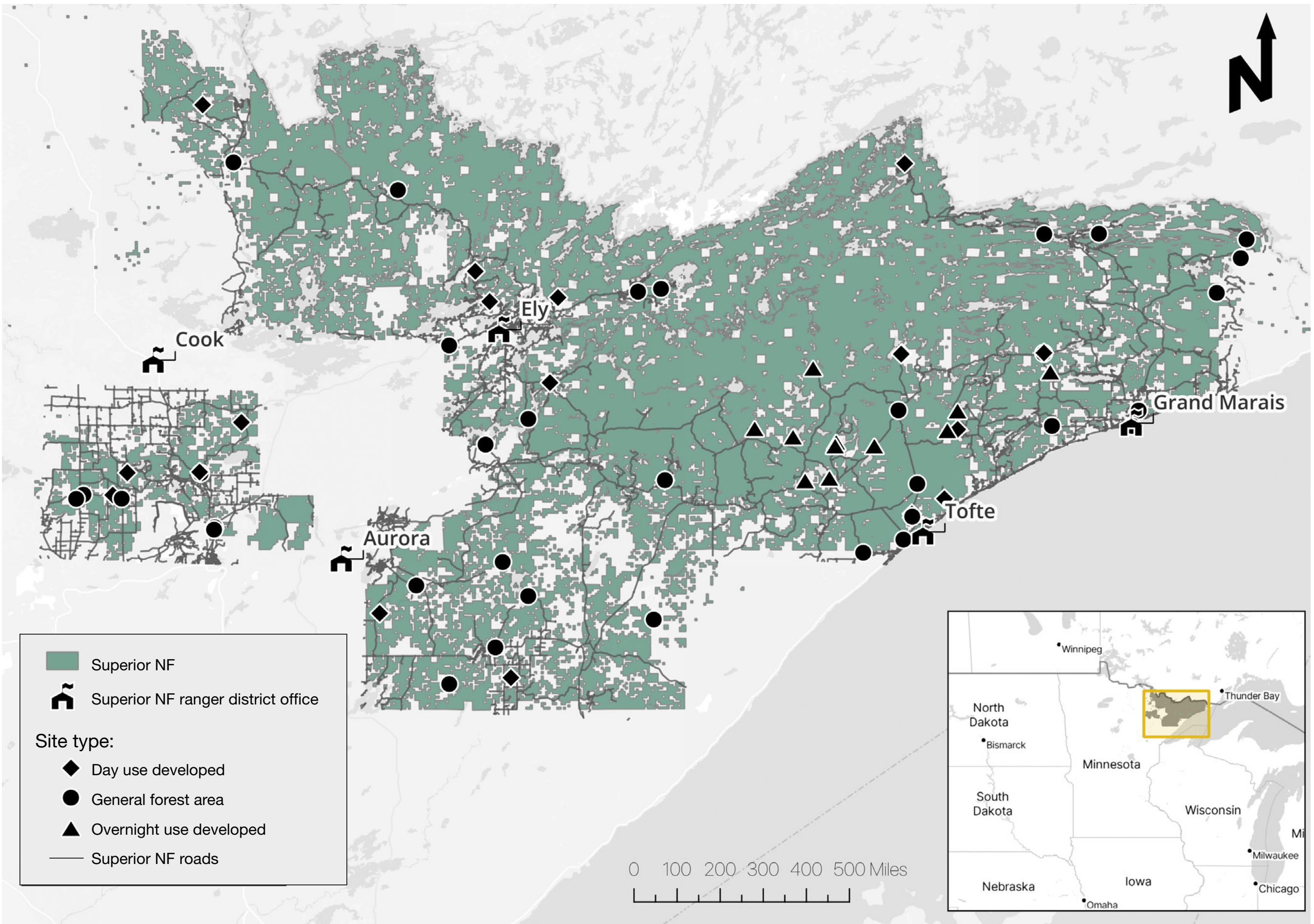

Figure 2—Superior National Forest area map with ranger district offices, site types, and OpenStreetMap drivable roads (OpenStreetMap Foundation, n.d.) used for travel cost calculation. NF = National Forest.

Optimization Problem Formulation

We implemented this MOO analysis using the pymoo (Blank and Deb 2020) Python library, with the binary inclusion or exclusion of each site-day (hexId) from the sample frame as the decision variable. The optimization sought to simultaneously minimize the three objective functions (travel cost, visitation pattern consistency, and sample representativeness), subject to two constraints (minimum number of site-days and target accuracy). We executed the optimization using the Non-dominated Sorting Genetic Algorithm II (NSGA-II), a widely used genetic or evolutionary algorithm for solving multi-objective problems inspired by transmission of favorable genes in natural selection in which candidate solutions from one generation are inherited by the subsequent generation (Yang 2021). This algorithm is highly effective at both exploring a wide search space and efficiently converging on a set of Pareto-optimal solutions. From this Pareto-optimal set, we selected a single, best-balanced scenario by identifying the solution with the lowest sum of the normalized objective function values.

$$\underset{x \in \{0.1\}^N}{\text{Minimize:}} \quad \mathrm{F}(x) = [f_1(x), f_2(x), f_3(x)]$$

$$\text{Subject to:} \quad \mathrm{g}_1(x) = n_s(x) \geq d_{min,s}, \ \forall_s \in S$$
$$\mathrm{g}_2(x) = \mathrm{Corr}(x) - 0.8 \geq 0$$

Where:

- x is a binary decision vector of dimension N (the total number of site-days in the full sample frame).
- ∈{0,1} is an element that indicates whether site-day i is included (1) or excluded (0) from the reduced sample.
- $F(x)$ is the vector of objective functions to be minimized:
  - $f_1(x)$ is the total travel cost for the selected site-days.
  - $f_2(x)$ is the MAD of last-exiting counts between the full and reduced samples across strata.
  - $f_3(x)$ is the Jaccard distance between the activity-type distributions of the full and reduced samples.
- $g_1(x)$ ensures that the number of site-days selected in each stratum meets or exceeds the minimum required number:
  - $g_s(x)$ is the number of site-days selected in stratum s.
  - $g_{min,s}$ is the minimum required number of site-days for stratum s.
  - S is the set of all sampling strata.
  - $\forall_s \in S$ indicates that this constraint applies to all strata.
- $g_2(x)$ ensures that the Pearson correlation coefficient between the mean last-exiting counts of the full and reduced samples is at least 0.8:
  - Corr(x) is the Pearson correlation coefficient comparing the sample's mean last-exiting counts to those of the full dataset across strata.
  - Corr(x)-0.8≥0 ensures that the correlation is at least 0.8.

We configured the optimization to run for 1,000 generations with a population size of 10, which was sufficient for convergence on a set of Pareto-optimal solutions given the small number of decision variables. Next, we specified the genetic algorithm parameters of **crossover** and **mutation**. Crossover simulates the mixing of chromosomes from parent genes, such that the offspring are a mix of the parent solutions, and provides a mechanism for convergence within a subspace (Yang 2021). On the other hand, mutation introduces random changes to the solutions, which increases the diversity of the population and provides a mechanism for the population to escape from local optima. To balance exploration and exploitation in the search space, we specified recommended values for the crossover probability (0.7), and mutation probability (0.2) (Hassanat et al. 2019).

We tasked the optimization to provide solutions for two sample-reduction scenarios: a 20-percent reduction in site-days per stratum and a targeted reduction scheme that allowed for fewer than eight site-days for one stratum (overnight use developed site [OUDS]-medium) (see table 1). After the termination criteria for the optimization are satisfied, either by meeting the maximum number of generations or by converging on a set of Pareto-optimal solutions, the optimization returns 10 sets of solutions; each solution is a potential sample frame that indicates the inclusion or exclusion of each site-day from the full sample frame.

Table 1—Number of site-days by site type and stratum for the full fiscal year 2016 Superior National Forest National Visitor Use Monitoring sample, the 20-percent reduction scenario, and the targeted reduction scenario

Site type	Use level	Full sample	20-percent reduction site-days	Targeted reduction site-days
		NUMBER	NUMBER (PERCENTAGE OF FULL SAMPLE)	
DUDS	Low	14	11 (78.6)	11 (78.6)
	Medium	14	11 (78.6)	12 (85.7)
	High	14	11 (78.6)	11 (78.6)
	Very high	9	8 (88.9)	9 (100.0)
GFA	Low	33	26 (78.8)	26 (78.8)
	Medium	19	15 (78.9)	19 (100.0)
	High	15	12 (80.0)	15 (100.0)
	Very high	15	12 (80.0)	15 (100.0)
OUDS	Low	14	11 (78.6)	14 (100.0)
	Medium	9	8 (88.9)	7 (77.8)
	High	8	8 (100.0)	8 (100.0)

Note: Across sample strata, we reduced the number of site-days by about 20 percent, while maintaining a minimum of 8 days in each stratum. In the targeted reduction scenario, we relaxed this minimum of 8 days for OUDS-medium.

DUDS = day use developed site, GFA = general forest area, OUDS = overnight use developed site.

Generalization via Classification Model

Because of the limited number of NVUM sampling sites common to fiscal years 2016 and 2026, we developed a binary classification model to predict the optimization results from fiscal year 2016 and extended them to the new and unobserved sites in fiscal year 2026. We developed the model using the binary decision variable for inclusion or exclusion of sites from the fiscal year 2016 optimization analysis as the target variable. Predictors for this model included the site geographic information (i.e., latitude and longitude), site type, use level (i.e., low, medium, high, or very high), expectations for visitation (i.e., number of no use, low, medium, high, or very high days), ranger district, scheduled survey time (i.e., am or pm), and temporal features (i.e., day of week, month, and day of year). Additionally, using OSMnx (Boeing 2025), we calculated the travel distance to the six ranger district office cities and towns for the Superior National Forest (Aurora, Cook, Duluth, Ely, Grand Marais, and Tofte). To more accurately represent the cyclical nature of the calendar year, we encoded the day of year feature with a radial basis function using scikit- lego (Warmerdam et al. 2025), which creates a smooth, continuous representation of time and enables the classification model to form simpler, more generalizable decision boundaries. We used a grid-search optimization to determine the optimal number of radial basis function periods to sufficiently but parsimoniously capture the seasonal patterns of the sample calendar. Next, we used a recursive feature elimination process to identify and remove predictors with low importance to the model's prediction performance (Kuhn and Johnson 2013). Finally, we used Optuna (Akiba et al. 2019) to tune the model hyperparameters minimizing the log-loss on 80 percent of the fiscal year 2016 data and validated the model on the remaining 20 percent of the data.

Results

Optimization Analysis

The MOO provided unique solutions for the two reduction in site-day scenarios: a 20-percent reduction across all sample strata and a targeted reduction to specific sample strata (table 2). For both scenarios, we identified the optimal solution as the candidate with the lowest sum of normalized values across the three objective functions. Figure 3 illustrates that the sampling cost (f_1) is linear, increasing with the number of site-days, while the trend lines for visitation pattern consistency (f_2) and sample representativeness (f_3) decrease with the number of site-days.

The 20-percent reduction scenario resulted in fewer total site-days (133) compared to the targeted reduction scenario (148) and relied on fewer unique sites (75) compared to the targeted reduction scenario (77). The accuracy and representativeness scores for the 20-percent reduction and targeted reduction scenarios were ideal, with minimal loss to visitation pattern consistency (MAD: 0.01 and 0.00, respectively) and sample representativeness (Jaccard distance: 0.06 and 0.02, respectively) (see Box 1 for the interpretation of optimization metrics). Because of the fewer site-days in the 20-percent reduction scenario, the travel-distance costs were lower ($3,907) compared to the targeted reduction scenario ($4,433). We estimated the sampling cost of data collection based on contract agreements in the Forest Service, Eastern Region (home to Superior National Forest) to be $384 per site-day. Accordingly, with fewer site-days, the fixed costs of sampling were lower for the 20-percent reduction scenario ($51,072) compared to the targeted reduction scenario ($56,832). Both reduction scenarios resulted in significant reductions in total costs compared to the full sample, with $15,582 (22.1 percent) in cost savings for the 20-percent reduction scenario and $9,295 (13.2 percent) in cost savings for the targeted reduction scenario.

Table 2—Summary metrics and costs of NVUM data collection for the fiscal year 2016 sample compared to the 20-percent reduction and targeted reduction scenarios identified using the multi-objective optimization framework

Metric	Full sample	20-percent reduction	Targeted reduction
Total site-days	169	133	148
Unique sites	84	75	77
Visitation pattern consistency (median absolute deviation)	0.000	0.010	0.000
Representativeness (Jaccard distance)	0.000	0.058	0.019
Travel cost	$7,584.41	$ 3,906.77	$ 4,433.09
Sampling cost	$62,976.00	$ 51,072.00	$ 56,832.00
Total cost	$70,560.41	$ 54,978.77	$ 61,265.09

Note: We calculated the fixed-sampling cost based on the average cost per site-day for National Visitor Use Monitoring (NVUM) data collection in the Forest Service's Eastern Region ($384/site-day); we calculated the travel cost as the sum of the travel-distance cost for each site in the sample frame for each scenario. All dollar amounts are in 2025 U.S. dollars.

Box 1—Interpretation of Optimization Metrics

- **Median absolute deviation (MAD)**: A robust measure of variability or spread in a set of numerical data because it is not affected by outliers. MAD represents the typical absolute distance between each data point and the median of the entire dataset.
- **Jaccard distance**: A metric used to measure the dissimilarity between two finite sets. It is the complement of the Jaccard index (which measures similarity). The Jaccard distance ranges from 0 to 1; a value of 0 means the two sets are identical (they contain all the same elements), and a value of 1 means the two sets are completely different (they have no elements in common).

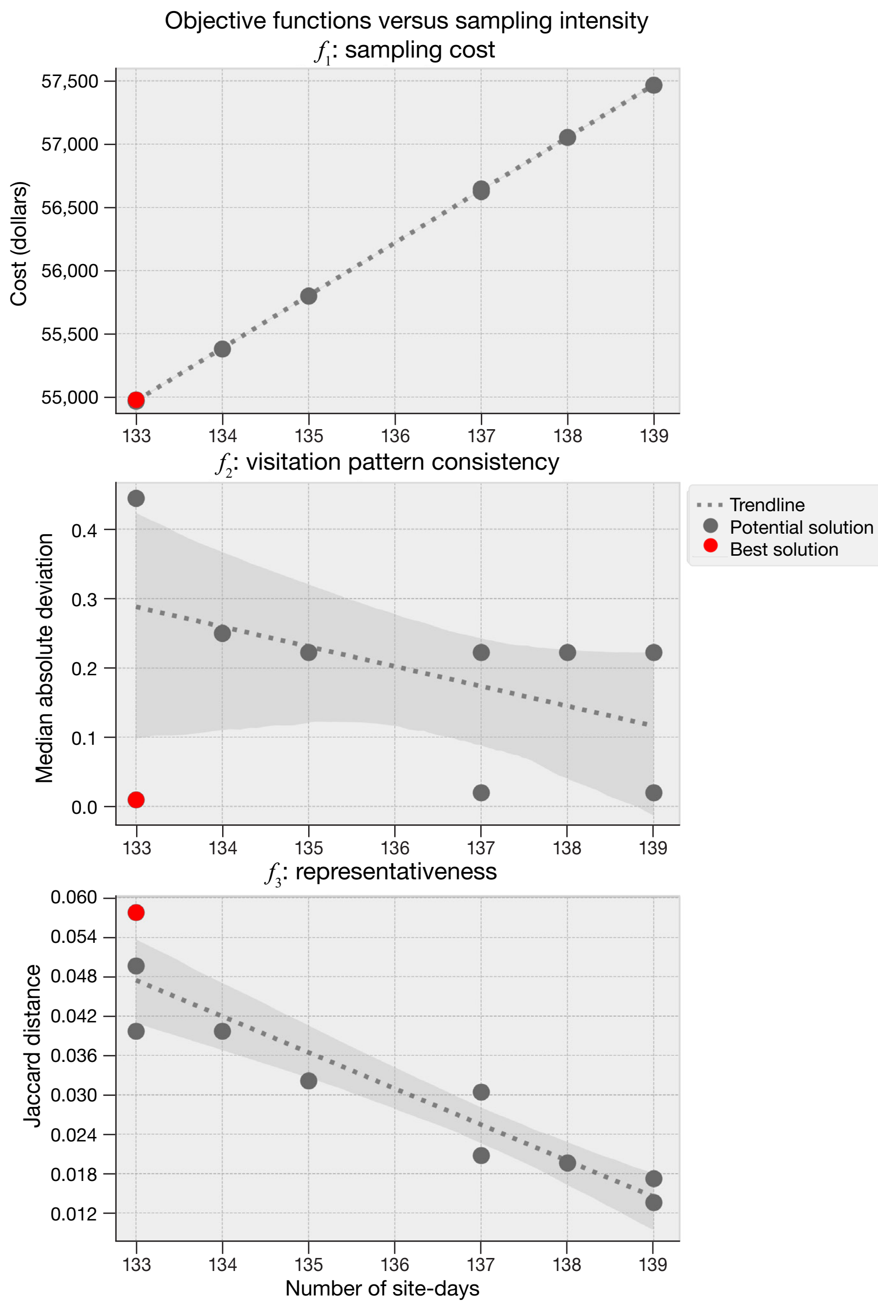

Figure 3—The solution set of optimizations for the 20-percent reduction in site-days scenario with objective functions plotted against the number of site-days in the solution. The best solution (red dot) was identified as the solution with the lowest sum of normalized values across the three objective functions. The dashed lines represent the trends for the objective as a function of the number of site-days.

Classification Model Projecting Optimization Solutions

The classification models for the two reduction scenarios, (20-percent reduction and targeted reduction), performed reasonably well on the fiscal year 2016 training dataset, with an accuracy of 0.853 and 0.912, respectively (table 3). Based on the F1-score, which provides a balanced measure of the classification model performance, the targeted reduction model (0.952) was more accurate than the 20-percent reduction model (0.906). However, the 20-percent reduction model showed stronger performance according to the Area Under Curve-Receiver Operating Characteristic score, 0.802 for the 20-percent reduction model compared to 0.625 for the targeted reduction model (see Box 2 for the interpretation of classification model metrics). We developed both models using the same methods and predictors, but we suspect that the simpler sample-reduction scheme for the targeted reduction model was more consistent and easier for the model to learn from the data.

Table 3—Classification model performance metrics for the 20-percent reduction and targeted reduction scenarios on the validation dataset for the fiscal year 2016 optimization of National Visitor Use Monitoring data collection

Full sample	20-percent reduction	Targeted reduction
Accuracy	0.853	0.912
Precision	0.923	0.909
Recall	0.889	1.000
F1-score	0.906	0.952
AUC-ROC	0.802	0.625

AUC-ROC = Area Under Curve-Receiver Operating Characteristic. See Box 1 for descriptions of terms.

Box 2—Classification Model Metrics

Metric	Description	Possible range
Accuracy	The percentage of correct predictions that the model made. Among all site-days, what fraction did the model correctly label as either “Retain” or “Drop.”	0.0–1.0
Precision	The ratio of correctly predicted positive observations to the total predicted positive observations. High precision score indicates a low false positive rate.	0.0–1.0
Recall	The ratio of correctly predicted positive observations to all positive observations. A high recall score indicates a low false negative rate.	0.0–1.0
F1-score	The balanced score (i.e., harmonic mean) of precision and recall. Higher scores indicate better performance.	0.0–1.0
AUC-ROC	Measures the model’s ability to discriminate between the two classes (“retain” or “drop”). A score of 1.0 means it can perfectly separate the “retain” days from the “drop” days, while a score of 0.5 means the model is no better than a random guess.	0.0–1.0

AUC-ROC = Area Under Curve-Receiver Operating Characteristic.

In figure 4, a confusion matrix illustrates the performance of each classifier at projecting the result of the optimization for the respective scenario. The confusion matrix provides a detailed breakdown of projection outcomes by comparing model projections against the actual or true values from the optimized sampling schedules. Each matrix is composed of four quadrants: true positives (top left) and true negatives (bottom right), which represent all correct classifications of retain or drop; false positives (top right), or type I errors, which occur when the model incorrectly projects retain (a false alarm); and false negatives (bottom left), or type II errors, which occur when the model incorrectly projects the site-days to be dropped (a miss). This granular breakdown is useful to move beyond single metrics to analyze and diagnose error between models.

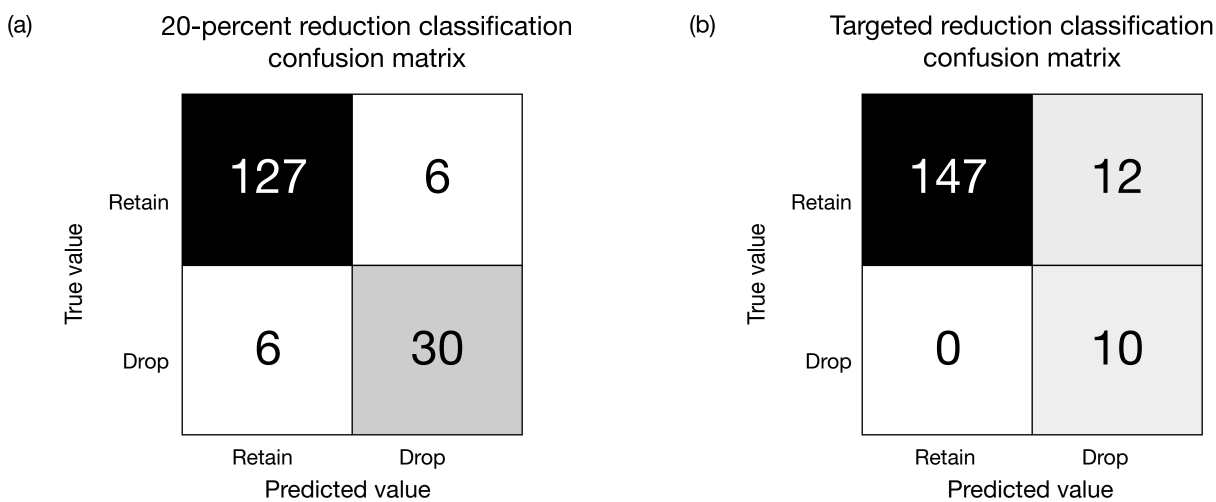

Figure 4—Confusion matrices for the 20-percent reduction and targeted reduction scenarios diagnosing classification model performance.

The 20-percent reduction classification model (fig. 4a) demonstrated strong performance, achieving an overall accuracy of 92.9 percent (157 correct classifications out of 169). The model demonstrated a well-balanced approach with a **precision** and **recall** of 95.5 percent. This indicates that the model was highly effective at identifying 127 of the 133 site-days to be retained, successfully retaining 127 of the 133 days in the sample. However, the model incorrectly projected 6 site-days to be dropped (false positives) and 6 site-days to be retained (false negatives).

The targeted reduction classification model (fig. 4b) also demonstrated strong performance, achieving an overall accuracy of 92.9 percent (157 correct classifications out of 169). The model had perfect precision (100.0 percent) in the retain projections, as zero site-days to be dropped were incorrectly retained (false positives). Further, the model had a high recall of 92.5 percent, successfully identifying 147 of the 159 total site-days to be retained, although the model incorrectly projected 12 days to be dropped that should have been retained (false negatives).

Discussion

Here we demonstrate a data-driven and systematic approach to reducing NVUM data collection costs while minimizing the deviation from established use estimates and representation of user activities relative to completing the full sample schedule. First in this two-part approach, we used a MOO to design an NVUM sample calendar based on a previous round of data collection (fiscal year 2016). A strength of the optimization approach is that it is data driven and context specific and leverages the substantial quantity of existing NVUM data to guide decision making for modifications to NVUM sampling methodology that yield cost reduction.

Second, we proposed a modeling approach to generalize the optimization strategy from past NVUM round datasets to a future round. This step addresses the challenge of the variation of sites available and selected for sampling in a forest from one round to the next. If the sampling sites were consistent between rounds, the reduced calendar could be adapted for a future round. However, because sampling sites can change between NVUM rounds, the optimized sample calendar from a past round cannot be seamlessly applied to a subsequent round. Consequently, we used the optimized sample calendar to train a binary classification model to predict which site-days should be retained or could be eliminated for a future NVUM sample calendar.

Application to NVUM Sampling and Methodology

A notable strength of these proposed approaches is their flexibility and adaptability to the context and composition of site types in a forest and their ability to explore a range of cost-reduction goals or scenarios, while also quantifying the impact or loss in the fidelity of the data. Furthermore, the objectives of the optimization analysis can be selected and adjusted for a range of metrics, such as the use estimates or activity types used in this analysis. Metrics could also include spatial and temporal activity patterns, respondent age, and home location. The sophistication of the MOO is its ability to efficiently identify site-days that satisfy the objectives of the analysis, while also distinguishing site-days that could be eliminated from the sample calendar with minimal impact on end uses of the data. Nevertheless, the MOO only functions retroactively, providing an optimized result from an existing sample calendar and survey dataset, which in practice often cannot be seamlessly applied or transferred to future sampling rounds. Our approach to overcoming this limitation was to develop a classification model to "learn" the strategy for the sample reduction from the optimization result and apply it to a future NVUM sampling round.

The estimated cost reduction from the optimization of the Superior National Forest for fiscal year 2016 was between 13 and 20 percent of the total survey costs, depending on the reduction scenario. If applied on all forests in a given year, these savings would provide nontrivial reductions in operating costs. However, smaller sample sizes decrease the granularity of the dataset, reducing the statistical power and possibly limiting the ability to analyze specific groups of forest visitors. A key strength of the NVUM program is the consistent sampling methodology that has been used across the sampling rounds. Modifying or reducing the intensity of NVUM sampling can reduce the ability to track changes in forest visitor characteristics, spending patterns, and evaluations of the forest experience over time and can reduce the utility of the data for evidence-based decision making and forest planning.

Applications for Long-term Monitoring Programs

Although the focus of this research note is to demonstrate data-driven, cost-reducing approaches within the context of the NVUM data-collection process, we anticipate that this methodological approach could be adapted for use in other long-term monitoring programs in which past data can be used to inform future data sampling decisions. The analysis used open-source Python packages for network analysis, MOO, and classification modeling, and was run on a standard laptop computer with no specialized hardware. While the code was written to be adapted for new forests, this approach could be readily adapted to other long-term monitoring programs by researchers or program managers interested in exploring similar approaches to systematically reduce sampling intensity. Large-scale survey data-collection programs are common in federal land management agencies, such as the U.S. Department of the Interior (USDI), National Park Service (Otak Inc. et al. 2024) and the USDI Fish and Wildlife Service (Dietsch et al. 2025), but mounting pressure from budgetary constraints increases the need for program managers to seek efficiencies in operations. The methodology outlined in this research note could be applied in other large-scale, longitudinal survey data-collection contexts to identify opportunities to reduce costs, increase the efficiency of data collection (e.g., via reduced travel costs), or even to empirically test the efficacy of cost-reduction strategies. Additionally, the optimization approach could be adapted to visitor-use monitoring applications aside from survey data collection, such as automatic trail or vehicle counters, to identify the most cost-effective monitoring strategy.

Limitations and Future Work

Despite the strengths of these proposed approaches, there are still limitations and challenges. First, conducting a robust evaluation of the quality of optimization and classification results on a subsequent sampling round remains unresolved. For the Superior National Forest, we lack a complete NVUM dataset from fiscal year 2021 to prescribe a reduced sample calendar and evaluate the impact of the reduced sampling intensity on the quality of the data. Ideally, this approach could be replicated on a unit with complete, high-quality datasets for three successive sampling rounds, with round one serving as a training dataset, round two serving as a validation dataset, and round three serving as a test dataset. However, this would still require the completion of the test data-collection round with the full sample calendar to definitively quantify the tradeoffs of this approach.

Second, this methodology, applying an optimization to a past round of data to inform the future round, makes the implicit assumption that the composition of the forest visitor population is stationary, along with visitor activity-type preferences and spatial and temporal patterns of use. Forests, as well as the visitors that use them, are dynamic and subject to change over time. As a result, we would expect demographic and socioeconomic changes and changes in preferences in recreation and activity type in the general population. We also expect the data collected to reflect behavioral responses to the increasingly pronounced effects of climate change. Although the 5-year time span between NVUM rounds is a relatively short time scale, we acknowledge that historical data may not be representative of future data, and its use may propagate bias into future data-collection rounds.

Future research could investigate the application of this optimization approach using multiple forests and rounds of NVUM data that could then be used to train and validate the classification model. Because the optimization can be customized to any given context (e.g., optimizing sample calendars for all forests within a region or in an NVUM sampling round), there are likely efficiencies and classification model performance benefits that could be realized by applying the optimization to multiple forests based on a longer temporal scope and with variations in the spatial characteristics of multiple rounds of NVUM data. A larger dataset for training the classification model could offer increased accuracy of predictions and efficiency benefits to apply to forests sampled in successive years without the need to optimize the sample calendar for each new forest.

Summary

Developing an approach to reduce NVUM data collection and sampling costs is complex and requires adaptability and flexibility that can account for variation from forest to forest and year to year. The application of MOO to NVUM sampling offers a potential approach to reducing the costs associated with NVUM data collection and allows program managers to make informed, data-driven decisions about where to focus limited resources. Here we demonstrated a two-part approach to addressing this problem; first, we used a MOO to identify an ideal set of site-days that minimizes the loss to sample representativeness and forest-wide use estimates. Second, to extend the sample reduction scheme to subsequent years, we developed a classification model to predict which days would be eliminated in the fiscal year 2026 calendar based on the results of the optimization of the fiscal year 2016 sample calendar. Nevertheless, the optimization approach outlined in this research note is not without limitations, and before any implementation, it is important to consider potential tradeoffs and uncertainties and more robustly validate and corroborate the results in a broader context.

References

Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. 2019. Optuna: a next-generation hyperparameter optimization framework. arXiv: 1907.10902. 10 p. https://doi.org/10.48550/ARXIV.1907.10902.

Askew, A.E.; English, D.B.K.; Zarnoch, S.J.; Poudyal, N.C.; Bowker, J.M. 2014. Mitigating budget constraints on visitation volume surveys: the case of U.S. national forests. Gen. Tech. Rep. SRS-GTR-192. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 21 p. https://doi.org/10.2737/SRS-GTR-192.

Blank, J.; Deb, K. 2020. Pymoo: Multi-objective optimization in Python. IEEE Access. 8: 89497–89509. https://doi.org/10.1109/ACCESS.2020.2990567.

Boeing, G. 2025. Modeling and analyzing urban networks and amenities with OSMnx. Geographical Analysis. 57(4): 567–577. https://doi.org/10.1111/gean.70009.

Dietsch, A.M.; Gutierrez, E.A.N.; Ducey, E.E.; Don Carlos, A.W.; Milley, B.J.; Sexton, N.R. 2025. Characterizing the experiences of visitors to national wildlife refuges: 2018–2023 national-level results. Columbus, OH: The Ohio State University, School of Environment and Natural Resources. 35 p. https://www.fws.gov/sites/default/files/documents/2025-09/nvs-2018-2023-national-report-final-508.pdf. (April 17, 2026).

Dumont, M.; Brunel, G.; Tresson, P.; Nespoulous, J.; Boukcim, H.; Ducousso, M.; Boivin, S.; Taugourdeau, O.; Tisseyre, B. 2024. Operational sampling designs for poorly accessible areas based on a multi-objective optimization method. Geoderma. 445: 116888. https://doi.org/10.1016/j.geoderma.2024.116888.

English, D.B.K.; Kocis, S.M.; Zarnoch, S.J.; Arnold, J.R. 2002. Forest Service National Visitor Use Monitoring process: research method documentation. Gen. Tech. Rep. SRS-GTR-57. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 14 p. https://doi.org/10.2737/SRS-GTR-57.

English, D.B.K.; Zarnoch, S.J.; White, E.M. 2018. Estimating daily exiting traffic from national forest recreation sites using short-term observations counts [Abstract]. In: Dehez, J., ed. The 9th international conference on monitoring and management of visitors in recreational and protected areas (MMV9). Bordeaux, France: IRSTEA Research Center (Environment, Territory and Infrastructure Research Unit). 252–254.

Hancock, J.M. 2014. Jaccard distance (Jaccard index, Jaccard similarity coefficient). In: Hancock, J.M.; Zvelebil, M.J., eds. Dictionary of bioinformatics and computational biology. 1st ed. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9780471650126.dob0956.

Hassanat, A.; Almohammadi, K.; Alkafaween, E.; Abunawas, E.; Hammouri, A.; Prasath, V.B.S. 2019. Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach. Information. 10(12): 390. https://doi.org/10.3390/info10120390.

Kuhn, M.; Johnson, K. 2013. Applied predictive modeling. New York: Springer. 443 p. https://doi.org/10.1007/978-1-4614-6849-3.

Li, X.; Gao, B.; Pan, Y.; Bai, Z.; Gao, Y.; Dong, S.; Li, S. 2022. Multi-objective optimization sampling based on Pareto optimality for soil mapping. Geoderma. 425: 116069. https://doi.org/10.1016/j.geoderma.2022.116069.

Lohr, S.L. 2021. Sampling: design and analysis. 3rd ed. New York: Chapman and Hall/CRC. 674 p. https://doi.org/10.1201/9780429298899.

Nayak, S. 2020. Multiobjective optimization. In: Fundamentals of optimization techniques with algorithms. London: Academic Press: 253–270. Chapter 9. https://doi.org/10.1016/B978-0-12-821126-7.00009-7.

Neyman, J. 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. Journal of the Royal Statistical Society. 97(4): 558–625. https://doi.org/10.2307/2342192.

OpenStreetMap Foundation. [N.d.]. Welcome to OpenStreetMap. https://www.openstreetmap.org. (July 30, 2026).

Otak Inc.; RRC Associates; University of Montana; GAIA Environmental Consulting. 2024. 2023 Socioeconomic research of National Park Service visitors: report on 2023 data collection. Fort Collins, CO: U.S. Department of the Interior, National Park Service, Natural Resource Stewardship and Science. https://irma.nps.gov/DataStore/DownloadFile/709436. (May 19, 2026).

Rashed, N.A.; Ali, Y.H.; Rashid, T.A.; Salih, A. 2024. Unraveling the versatility and impact of multi-objective optimization: algorithms, applications, and trends for solving complex real-world problems. Journal of Soft Computing and Computer Applications. 1(1): 1008. https://doi.org/10.70403/3008-1084.1008.

Tillé, Y.; Wilhelm, M. 2017. Probability sampling designs: principles for choice of design and balancing. Statistical Science. 32(2): 176–189. https://doi.org/10.1214/16-STS606.

U.S. General Services Administration. 2025. POV mileage rates (archived). https://www.gsa.gov/travel/plan-a-trip/transportation-airfare-rates-pov-rates/privately-owned-vehicle-pov-mileage-reimbursement/pov-mileage-rates-archived. (May 26, 2026).

Warmerdam, V.D.; Bruzzesi, F.; MBrouns; Collot, S.; de Boer, J.; Kübler, R.; pim-hoeven; Mkalimeri; Paulino, A.; Gorelli, M.E.; Verheijen, P.; Borry, M.; Hoogland, K.; Masip, D.; Kowalczuk, M.; ktiamur; AminaZ; Sharma, G.; Lepelaars, C.; Rens; Cor; Hermans, F.; Borrella, T.; Rose, S.A.; van Dorsten, S.; Levitski, H.; Keromnes, J.; Pérez-Lozao, S.C.; Payne, S. 2025. Zendo. Koaning/scikit-lego: v0.9.5. [Software]. https://doi.org/10.5281/ZENODO.15313097.

Yang, X.-S. 2021. Nature-inspired optimization algorithms. 2nd ed. San Diego, CA: Academic Press. 310 p.

Zarnoch, S.J.; English, D.B.K.; Kocis, S.M. 2005. An outdoor recreation use model with applications to evaluating survey estimators. Res. Paper. SRS-RP-37. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 15 p. https://doi.org/10.2737/SRS-RP-37.

Zarnoch, S.J.; Kocis, S.M.; Cordell, H.K.; English, D.B.K. 2002. A pilot sampling design for estimating outdoor recreation site visits on the national forests. Res. Paper. SRS-RP-29. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 20 p. https://doi.org/10.2737/SRS-RP-29.

Glossary

accuracy—the percentage of correct predictions that the model made.

Area Under Curve-Receiver Operating Characteristic (AUC-ROC)—a measure of how well a model performs at distinguishing between two classes. A score of 1.0 is perfect, while a score of 0.5 indicates a classifier would perform no better than a random guess.

constraints—the rules or restrictions that are applied to the optimization problem to ensure that the solution is feasible and meets the objectives.

crossover—a process used in genetic optimization algorithms in which two candidate solutions (parents) exchange and blend parts of their "genetic" information to create new solutions.

daily summary form—a dataset that summarizes daily interview results, including the calculated traffic count, number of visitor contacts, and number of last-exiting visitor interviews for each site-days.

F1-score—a measure of the accuracy of a classification model that is calculated as the harmonic mean of precision and recall.

Jaccard distance—a measure of the similarity between two sets of populations, calculated as the size of the intersection divided by the size of the union of the two sets.

last-exiting—a category that denotes visitors who recreated at the site and are leaving the site or forest for the last time that day.

median absolute deviation (MAD)—a measure of the spread of a dataset that is calculated as the median of the absolute deviations from the median of the dataset.

mutation—a process used in genetic optimization algorithms in which a small, random change is made to a candidate solution to introduce new traits and maintain diversity in the set of solutions that prevent the search from being constrained to local optima.

nonproxy site—recreation sites or areas where National Visitor Use Monitoring program technicians interview recreation visitors to gather information about their visit and estimate recreation visitation.

objective functions—the goal of an optimization problem, which can either be minimized or maximized based on the problem at hand.

precision—the percentage of predicted positive instances that were true positive instances.

proxy site—recreation sites or areas where visit information from permits, fee envelopes, or ski lift tickets can be used to quantify the amount of recreation visitation.

recall—the percentage of true positive instances that were predicted as positive instances.

sample strata—the combination of site type and use level (e.g., DUDS-high, OUDS-low) for nonproxy sites. See "site type."

site-day—the interview location and date when National Visitor Use Monitoring program sampling occurs.

site type—one of four types of recreation sites defined in National Visitor Use Monitoring sampling: day use developed site (DUDS), overnight use developed site (OUDS), general forest area (GFA), designated wilderness (wilderness).

use level—the expected amount of last-exiting people or vehicles at a given interview location that are surveyed during National Visitor Use Monitoring program sampling. The five use levels are low, medium, high, very high, and none (closed).

view corridor site—sites intended to measure the number of visitors that only pass through or close to the forest to view the scenery while traveling on highways owned by the state, county, or federal transportation authority.

Pacific Northwest Research Station

Website	https://research.fs.usda.gov/pnw
Telephone	(503) 808-2100
Publication requests	(503) 808-2138
FAX	(503) 808-2130
E-mail	sm.fs.pnw_pnwpubs@usda.gov
Mailing address	Publications Distribution Pacific Northwest Research Station USDA Forest Service 1220 SW 3rd Avenue Portland, OR 97204