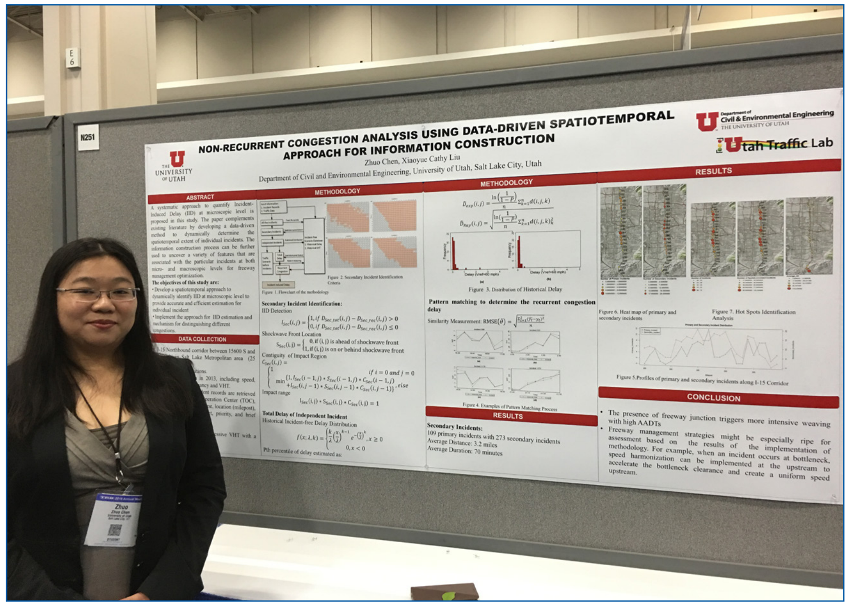
MOUNTAIN-PLAINS CONSORTIUM

MPC 18-350 | C. Liu and Z. Chen

Statistical Analysis and Sampling Standards for Maintenance Management Quality Assurance (MMQA)



MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

STATISTICAL ANALYSIS AND SAMPLING STANDARDS FOR MAINTENANCE MANAGEMENT QUALITY ASSURANCE (MMQA)

Xiaoyue Cathy Liu, Ph.D., P.E.
Assistant Professor
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Phone: (801) 587-8858
Email: cathy.liu@utah.edu

Zhuo Chen, Ph.D.
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Phone: (801)-541-1927
Email: zhuo.chen@utah.edu

June 2018

Acknowledgements

The authors acknowledge the Mountain-Plains Consortium (MPC) and the Utah Department of Transportation (UDOT) for funding this research, and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Rukhsana Lindsey
- Lloyd Neeley
- Tim Ularich
- Jason Richins
- Tammy Misarasi
- Todd Richins

ABSTRACT

Maintenance management has relied heavily on collecting asset condition information to plan maintenance activities and budget allocation. Data collection is often conducted on a sampling basis because of resource constraints. There is thus a perceived need for the development of an effective sampling framework that can determine statistically representative samples, reflect the true level of maintenance (LOM) at state, region, and station levels, and accommodate agencies' requirements. This project advances existing knowledge by presenting a systemic approach for a sampling scheme development to assist maintenance activity planning. The proposed method addresses how much and where agencies need to collect asset condition data for accurate LOM estimation. The method integrates Fisher information with a spatial sampling technique that can be customized based on local agencies' requirements, such as station balanced, spatially balanced, or others. The framework is showcased via an example application of the Signage Repair and Replace database maintained by the Utah Department of Transportation (UDOT). Four sampling methods that might be tempered to various needs are implemented. Sampling results are presented and compared against historical full asset inventory via similarity analysis. The proposed framework lays a strong theoretical foundation for maintenance asset sampling and is effective for estimating LOM at state, region, and station levels to assist with budget allocation. The method can be easily transferable and adoptable to other agencies for optimal maintenance management.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Maintenance management has been a critical component in today's transportation system, since a sustainable network relies heavily on the preservation of its infrastructure assets. An optimal maintenance program has always been focused on refining maintenance activities such that limited resources can be effectively used. The ability to report maintenance conditions with satisfying accuracy and efficiency is largely determined by data availability. Yet collecting such information can be demanding in terms of the resources, personnel, and time required. It is thus desirable to collect asset conditions on a sampling basis rather than on the entire asset inventory without loss of fidelity for unveiling the true asset conditions. Pertaining to maintenance management, the primary goal of sampling is to successfully estimate the overall LOM at state, region, and station levels to assist with budgeting and resource allocation. To fill the gap, this project presents a systematic approach for developing a sampling scheme customized to maintenance activities. The proposed method addresses how much and where agencies need to collect asset data with the maximum information retained for LOM estimation. The method integrates Fisher information with a spatial sampling technique that can be customized based on local agencies' requirements, such as station balanced, spatially balanced, or functional class based. These requirements are rooted in the very fundamentals of maintenance management.

The proposed sampling framework was showcased via an example application of the Signage Repair and Replace database maintained by the UDOT. The sampling method was enhanced on the basis of generalized random-tessellation stratified (GRTS) design by tailoring it to the maintenance setting. Four sampling methods that might be tempered to various needs were implemented, including GRTS sampling with equal segment weight, GRTS sampling weighted by signage, spatially-balanced sampling with equal segment weight, and spatially-balanced sampling weighted by signage. The sampling results were presented and compared against ground truth asset inventory. Comparing with simple random sampling method that is widely used by agencies across the country, the GRTS sampling with equal segment weight and the spatially balanced with equal segment weight methods both demonstrate better performance with much lower average similarity score. It is shown that the proposed framework lays a strong theoretical foundation for the maintenance asset sampling, based on the customized requirements/needs for local agencies, and is effective in estimating LOM at state, region, and station levels for budget allocation.

1. INTRODUCTION

1.1 Problem Statement

A sustainable transportation system relies heavily on the preservation and maintenance of highway assets to ensure and improve its functionality. The adoption of performance-based transportation management has been gaining popularity as the key feature in the Moving Ahead for Progress in the 21st Century Act (MAP-21). It further motivates the need for a streamlined process to make transportation investment decisions on the basis of asset performance. Accurate reporting of asset condition is critical to maintenance planning as it helps identify where and when assets must be reconstructed or replaced and consequently drives budget allocation and project prioritization. The impact of maintenance activities on such has led to the design and implementation of numerous initiatives to improve maintenance quality and establishment of quality assurance programs nationwide (Yurek et al., 2012). Over the years, these programs have evolved to focus more on effectively reporting maintenance outcome and reaching the targeted level-of-maintenance (LOM).

The key point getting great attention is the ability to report maintenance conditions with satisfying accuracy and efficiency, which in large is determined by data availability. Of further interest, data reduction is an indispensable component of today's transportation management. Inspecting highway assets can be demanding in terms of resources, personnel, and time required. It is desirable to conduct inspection on a sampling basis rather than on the entire asset inventory, yet with a loss of fidelity that is negligible in determining the true asset conditions. Therefore, developing proper sampling techniques to manage a region's assets and accurately infer the LOM based on statistically representative dataset has become an intriguing topic over the past decades. Most state agencies, such as Florida Department of Transportation (DOT), Indiana DOT, and Colorado DOT—to name a few—use simple random sampling (SRS) method in their maintenance quality assurance programs to introduce randomness into the sampling process, yet they lack consideration on assets' spatial distribution and justification on the representativeness of the sampled data (Schmitt et al., 2006). Challenges lie in the inherently missing approach that is theoretically sound for accurately choosing asset samples reflecting the LOM for decision making and budget allocation. Biased sampling can further cause well-intentioned policies to produce unintended consequences.

1.2 Objectives

The overarching goal of this research is to accurately estimate statewide LOM distribution on the basis of statistically selected samples. This is an important metric for many maintenance applications. The primary objective of this project is to develop systematic approach to assess "where and to what extent" to collect asset conditions with maximum information retained for LOM estimation. The challenges in developing such approaches vary, as the types of assets change.

The accuracy and efficiency of sampling method are closely associated with the sample size and the spatial correlation between sampled segments. The sampling method determines sample size based on data-driven analytics rather than intuitively. Fisher information is calculated to estimate the minimum sampling rate sufficient to capture the asset conditions throughout the network. Although it has been intensively applied in the realm of statistical modeling for its effectiveness in measuring information contained in samples, Fisher information has never been used in transportation for asset management and optimization. To select spatially well-distributed samples, rules combining Generalized Random-Tessellation Stratified Design (GRTS) and hierarchical randomization are applied in the sampling framework. This spatially balanced sampling technique is capable of accommodating customized needs of local agencies to ensure that collected assets are representative across a defined spatial coverage. The proposed method is a useful contribution easily adoptable to many agencies for optimal maintenance management.

1.3 Scope

Tasks in developing our proposed sampling method include:

- Construct Fisher Information with asset conditions to estimate the minimum sampling rate that is sufficient to capture the asset LOM in the network; and
- Develop an algorithm combining GRTS and hierarchical randomization to select spatially-balanced sample.

To conduct the above-mentioned tasks requires the support of a massive amount of historical data from multiple sources and jurisdictions. The asset inspection records are provided by the Utah Maintenance Management Quality Assurance (MMQA), a program established by the UDOT in 1997 for evaluating and reporting the effectiveness of its maintenance activities. The program has evolved to provide systematic guidance on feature condition thresholds, funding projection and allocation, and LOM measurements. MMQA offers guidelines on 17 measurement activities such as snow and ice, litter pickup, vegetation control, etc. It further refines specifications on the criteria of desired/defect conditions of each activity/asset. Inspectors are required to be familiar with the procedure and methodologies described for each maintenance activity before going into the field. The graphical description in MMQA helps them confidently describe the condition of any particular asset. Maintenance performance is measured and reported in the form of LOM, expressed as 15 different letter grades (A+ to F-). The entire statewide highway system is divided by 76 maintenance stations. Each station further divides each of its routes into one or more segments (2,048 segments in total). The personnel conduct inspections for each route segment, and record the total number of assets to be maintained on that segment and the total number of defect assets.

1.4 Outline of Report

The remainder of this report is structured as follows. Section 2 summarizes literature on maintenance asset condition sampling and spatial sampling techniques. The proposed sampling approach and its mechanism are presented in Section 3. Section 4 describes the data sources used for testing the sampling methods and Section 5 presents the results via an example application of the *Signage Repair and Replace* database maintained by the Utah Department of Transportation (UDOT). Conclusions and implications are discussed in Section 6.

2. LITERATURE REVIEWS

2.1 Sampling Techniques in Maintenance Management

Many state DOTs have developed maintenance quality assurance (MQA) program guidelines, most of which adopt certain forms of simple random sampling techniques for asset data collection (Schmitt et al. 2006). Simple random sampling chooses segments randomly by applying a fixed sampling rate. The probability of each segment being chosen is the same. With simple random sampling, network segmentation directly affects sampling efficiency. For a given network, the sample population is determined by the length of sample segment, as maintenance activities are conducted segment by segment. A long segment leads to a small sample population and consequently increases the sampling rate to meet the requirement of minimum sample size. However, a short segment leads to an increase in labor hours for collecting the data, since the maintenance personnel may need to drive through more unsampled segments between sampled segments. The selection of segment length is an empirical process—a decision made by maintenance operators. For example, the California Department of Transportation (Caltrans) and New York Department of Transportation (NYDOT) use 1 mile as a sampling unit; North Carolina uses 0.2 miles; and Florida, Indiana, Texas, Virginia, Washington, and Wisconsin use 0.1 mile (Schmitt et al. 2006). Once the sampling segment unit is determined, the question is directed to the selection of the sample size.

Three methods have been widely used in previous studies to determine sample size: fixed percentage of population (Templeton and Lytton 1984), the statistical method (De la Garza et al. 2008; McCullouch and Sinha 2003; Medina et al. 2009; Schmitt et al. 2006; Selezneva et al. 2004), and the optimization method (Gharaibeh et al. 2010; Mishalani and Gong 2008, 2009). Among the three, fixed percentage of population is easy to implement, yet accuracy is compromised because of its empirical nature and a lack of scientific validation. The proportion of samples needed from the entire population varies by the type of maintenance and rehabilitation (M&R) activities and needs. Templeton and Lytton (1984) believed that a sample size of 30–35% is needed to predict the cost to repair segments below a certain condition threshold. Among the surveyed transportation agencies that have MQA programs, maintenance sampling range varied from 1.5% to 20% (Yurek et al., 2012). Statistical methods are based on approximated sampling distributions, appearing to be more statistically valid compared with a fixed percentage of population. Schmitt et al. (2006) summarized a series of applications of standard statistical methods in maintenance sampling, such as using confidence interval of normal population, number of observations for t test of mean, etc. Selezneva et al. (2004) tested sample sizes on different randomly picked reliability levels until the corresponding sample met the requirement of quality assurance criteria. One novel method in determining sample size is by optimizing the maintenance plan. Most studies using optimization techniques apply Latent Markov Decision Process (LMDP). LMDP is a classic approach to solve long-term network-level M&R policy optimization problem. The purpose of LMDP is to maximize performance (e.g. higher LOM) with a given budget or to minimize costs with required performance. For example, in pavement maintenance, LMDP can determine how to assign routine maintenance, resurfacing, and inspection activities to the network (Mishalani and Gong 2008). Mishalani and Gong (2009) considered sample size as a decision variable in the LMDP optimization framework. Research on LMDP in terms of sampling rate is limited by far. Compared with the two aforementioned methods, LMDP is used specifically for maintenance activities and flexible for different types of assets. Gharaibeh et al. (2010) optimized sample size by minimizing the costs of performing sampling and the equivalent cost of inconvenience caused by poor-quality materials and construction. Due to the complexity of optimization, simplified assumptions are often made for the probability function in optimization methods, compromising the model fidelity. To implement optimization methods, it also requires good historical database to ensure accurate construction of transition matrix.

Once sample size is determined, a sampling plan must be designed to obtain features of interest. Several sampling design schemes have been widely used, including simple random sampling, sampling with replacement, sampling without replacement, stratified random sampling, etc. For MQA, simple random sampling and stratified random sampling are the most popular methods (Schmitt et al. 2006). However, as pointed out by previous researchers, the accuracy of true population condition estimate not only depends on the quality of measurements and sample size, but also on correlation among asset conditions at different locations (Mishalani et al., 2011). Such spatial correlation exists in maintenance sampling. As Mishalani et al. (2011) mentioned, smaller spatial correlation leads to more accurate estimation of asset conditions. Simple random sampling and stratified random sampling do not take this into account.

2.2 Spatial Sampling

In the context of spatial sampling, samples are collected typically in 1-, 2- or 3-dimensional space. Generic situations arise when the resource population is represented as collections of points, lines, or areas over spatial extents. Spatial sampling can be conducted using the traditional sampling methods mentioned in the foregoing cited studies, such as simple random sampling, systematic sampling, stratified random sampling, just to name a few. It can also take into account the unique spatial features resided in population, such as spatial autocorrelation and spatial heterogeneity (Goodchild et al., 1992; Ripley, 2005; Wang et al., 2012). Previous studies have applied a variety of spatial sampling techniques that appear to perform reasonably well in different sampling applications for getting a spatially-balanced sample. Yet still, numerous difficulties present themselves in multiple occasions. For example, when applying stratified sampling on one- or two-dimensional populations, it is difficult to split the entire population into spatially contiguous strata, especially when variable probability or substantial variations in spatial density exist (Stevens and Olsen, 2004). Spatial stratification has a wide range of applications due to the fact that as heterogeneity can be reduced in stratum and the ease of collecting samples that are highly representative (Wang et al., 2010). Among them, the Generalized Random-Tessellation Stratified Design (GRTS) provides a flexible means for selecting spatially well-distributed samples (Stevens and Olsen, 2000). By combining the GRTS with hierarchical randomization, it maps data from two-dimensional space into a one-dimensional linear structure for sampling, which can eventually result in a spatially well-balanced random sample. The method is well-suited to be employed in the sampling procedure for maintenance activities, especially given the needs of transportation agencies in terms of maximized spatial coverage when collecting data.

2.3 Summary

This section summarized key findings from literature research for this study. Two major issues are reviewed with regard to asset sampling in this project including sample size and uneven spatially distributed sample. To address these issues, we apply Fish information and spatially-balanced sampling techniques to the maintenance assets. Previous studies applied these methods/algorithms mainly in statistics, electrical engineering, and computer science. In the following sections, we demonstrate how these methods can be used to determine the asset sampling for optimal maintenance management.

3. RESEARCH METHOD

3.1 Overview

This section describes our proposed sampling method for maintenance asset condition estimation. The proposed method integrates Fisher information with spatial sampling techniques and can accommodate local agency's needs (e.g. sample on various functional classes, stations, maximum spatial coverage, etc.). It is also flexible for potential integration with spatial optimization to set certain resource constraints. The proposed framework lays a strong theoretical foundation for maintenance asset sampling and is effective in estimating LOM at state/region/station levels for budget allocation.

3.2 Fisher Information for Determining Sample Size

Fisher information is a measure of information that is expected in a trial X about the parameter θ. It can be defined as the derivative of the log-likelihood function with respect to θ (Ly et al., 2014):

$$\mathrm{I}(\theta) = \mathrm{Var}(\frac{d}{d\theta} \log f(X|\theta)) = -\mathrm{E}(\frac{d^2}{d\theta^2} log f(X|\theta)) \tag{1}$$

Fisher information has been applied hin a variety of statistical paradigms to answer different substantive questions. Liu and Yu (2009) used Fisher information to determine optimal geolocation data compression ratio for transportation target identification. Towsley et al. (2006) applied Fisher information metric to determine flow size distribution from packet sampling for network monitoring. At its minimum, numerous other studies have used it to either define a prior default parameter, determine sample size, or measure model complexity (Lee and Wagenmakers, 2013; Myung, 2003; Rissanen, 1996; Stevens, 1957). It plays a pivot role in statistical modeling.

In the context of the maintenance asset sampling scheme, Fisher information is a measurement of the maximum likelihood that the sampled maintenance inspection outcome represents the true asset condition. It can be used to determine the appropriate sample size. Take signage inventory as an example. The inspection can be treated as a Bernoulli process, where the sign's condition is either desired ("1") or defect ("0"), or vice versa. The probability density function (pdf) of a Bernoulli model can then be expressed as:

$$\mathrm{f}(\mathrm{x}|\theta) = \mathrm{P}(\mathrm{X} = \mathrm{x}) = \theta^x (1-\theta)^{1-x}, \text{ where x=0 or x=1} \tag{2}$$

where θ is the probability that the sign's condition takes on the value of 1.

The Fisher information of sample from Bernoulli model can be calculated by plugging Equation (2) into Equation (1), which yields:

$$\mathrm{I}(\theta) = -\sum_{x=0}^{1} \frac{d^2}{d\theta^2} log P(X=x) P(X=x) = \frac{1}{\theta(1-\theta)} \tag{3}$$

As shown in Figure 3.1, Fisher information demonstrates the sensitivity of a Bernoulli model with respect to parameter θ. As Fisher information increases, the sample becomes more accurate in describing the real condition of the population. When θ reaches 0 or 1, the expected Fisher information goes to infinity. Namely, when the conditions of signs are all "desired" or "defect," any sample can perfectly reflect the real condition of all segments.

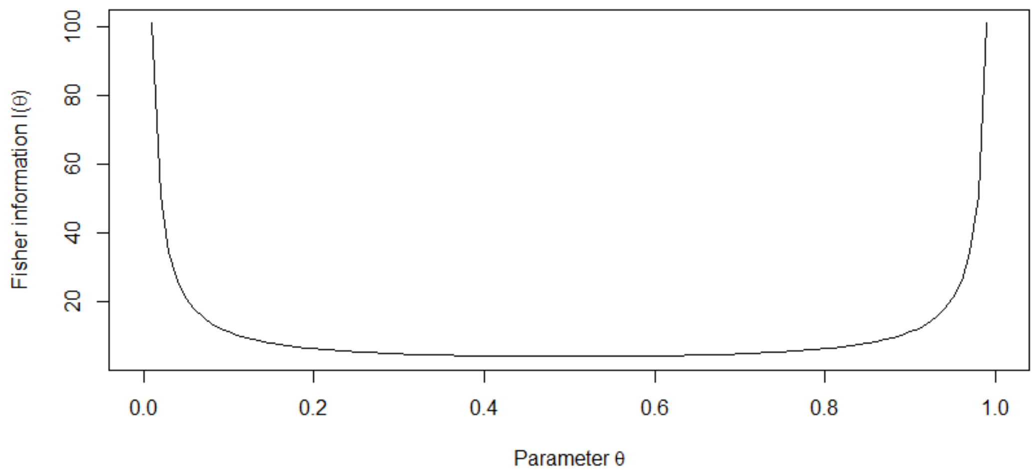


Figure 3.1 Fisher information as a function of θ in Bernoulli model

In reality, we only observe a single outcome $\vec{x}$ (the maintenance inspection at certain time) of size n (sample size) and have to infer θ instead. The goal for maintenance management is to provide a reasonable guess of the true value θ^*, such that the true LOM is unveiled. Fisher information can be used to determine the asymptotically least number of trials n that must be collected, such that an estimator $\vec{X}$ yields estimates $\vec{x}$ at a certain level of accuracy. For a complete derivation on measuring the performance of an estimator of a Bernoulli Model, interested readers are welcome to refer to (Ly et al., 2014). The general consensus is that as the number of trails n increases, more information is extracted about θ. With an independent and identically distributed (i.i.d.) assumption for $\vec{x}$, variance of the estimator $\vec{X}$ is given by: $\mathrm{Var}(X) = \theta(1-\theta)$. The goal is to tame the variance such that the largest variation is minimized, which occurs at $\theta = 0.5$ (as shown in Figure 3.1).

To determine the sample size n for the signage inspection $\vec{X}$ with $X \sim \mathrm{Ber}(\theta)$, the problem can be formulated such that the chance of obtaining an estimate that is more than α distance away from the true value is no larger than β, which can be expressed as:

$$P(\bar{X} \in [0.5 \pm \alpha]) \gg 1 - \beta \tag{4}$$

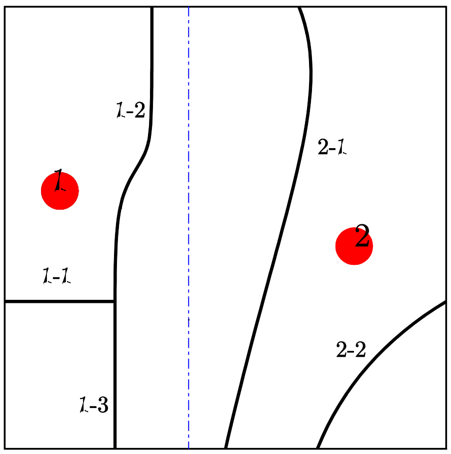
3.3 GRTS for Maintenance Management

For maintenance management purposes, sampling process should be able to accommodate varying spatial sample intensity, and spread the sample points evenly and regularly over the domain. Most agencies are using simple random sampling for maintenance management, yet it tends to exhibit uneven spatial patterns. Signage population, as an example, exists in spatial matrix. Although systematic sampling might compensate on the spatial feature, it has limited flexibility in changing sample point density or accommodate various inclusion probability. GRTS design combines simple random and systematic characteristics, and guarantees all possible samples are distributed across the resource. The basic idea of GRTS method is to create a quadrant-recursive function that maps two-dimensional space into a one-dimensional one, thereby defining an ordered spatial address for the population. Unequal probability sampling can be achieved by giving each point a length proportional to its inclusion probability.

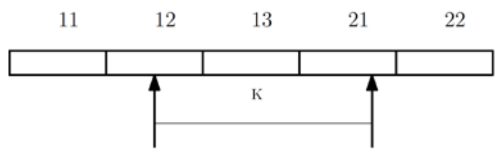
In maintenance application, the target population is the roadway segments partitioned by stations. To illustrate GRTS sampling scheme, Figure 3.2 shows an example region where five roadway segments are under the maintenance jurisdictions of two stations (circled). The segments are randomly labeled according to the station ID. Different from the classic quadrant partitioning of a region, the maintenance application already labeled (or spatially partitioned) each segment by stations. Therefore, when mapping the two-dimensional space into an ordered 1-dimensional linear structure, the features in Figure 3.2 (a) would be transformed into Figure 3.2 (b). Note that in Figure 3.2 (b), each segment is assigned equal probability, yet unequal probability can be tempered by the allowance of unequal length for each unit, as shown in Figure 3.2 (c). The sampling scheme can be expressed as:

$$d + (i - 1) \times k \qquad for\ i = 1,2, \ldots, n \tag{5}$$

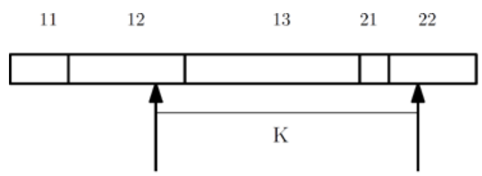
where *d* is a random start within the 1-dimensional space along [0,*k*), *L* is the total length of line, *n* is the sample size, and *k*=*L*/*n* .



(a) Example region for maintenance activities



(b) Linear frame with equal sampling probability



(c) Linear frame with unequal sampling probability weighted by the number of signs within each roadway segment

Figure 3.2 Example of GRTS algorithm for maintenance activity sampling

In Figure 3.2 (b)'s example, two out of five of the segments are sampled (n=2). If agencies desire segments that contain more signs to be sampled, then the length of the line can be entertained to represent signage amount within, as shown in Figure 3.2 (c), and the sampling result will vary correspondingly. Properly selected k can ensure that each station has at least one segment being sampled. This will be discussed in length in the example application.

3.4 Spatially-Balanced GRTS Sampling for Optimal Maintenance Management

The aforementioned method ensures the segments are ordered in the sequence of randomly labeled stations. And with properly chosen k, each station will have one or more segments selected. However, some agencies would prefer to have a spatially-balanced sample rather than a station-balanced sample when reporting LOM. To fulfill such spatially-balanced sampling feature, a hierarchical randomization can be applied to randomly order the generated addresses based on the quadrant-recursive function in GRTS.

In the classic GRTS design, a grid is divided into four cells, each of which is further divided into four subcells, and so on. The quadrant-recursive function is defined by the limit of the successive intensification in the grid. The recursion is carried through division and each division will pair the point with an address based on the order that the division was performed, where each digit of the address represents a step in the subdivision. A spatially referenced address can be constructed following the pattern of such partitioning. Random permutation that defines the hierarchical randomization is performed. Such recursive partitioning generates a nested hierarchy of grid, and puts the sampling process in the entire spatial context. Note that the order corresponds to the ranking obtained by reversing the sequence of the base-4 digits and treating the reversed sequence as a base-4 fraction. The reverse hierarchical order then gives a spatially well-balanced sample (Stevens and Olsen, 2004).

Different from the classic GRTS design that follows quadrant-recursive function with the resulting address appears as digits in a base-4 fraction, in a maintenance management setting, the segments are already partitioned in each station with varying sizes, leaving the difficulty of creating the address with a consistent base-N fraction. We remedy this with the following approach. Assume the maximum possible number of segments that a station has in the entire region is *N*, and the number of segments contained in station *i* is *n*, then the digit assigned to the ordered segment *j* can be expressed as:

$$T_{ij} = \begin{cases} 1, & when\ j = 1 \\ \frac{N}{j}, & otherwise \\ N, & when\ j = n \end{cases} \tag{6}$$

Reverse hierarchical order can again be applied to this base-N fraction via reversing the digits and sorting. The generated sequence will be reduced to the linear frame discussed in the previous section and is available for sampling with equal/unequal probability. The result will be a contiguous set of sample segments that are spatially-balanced, where the samples are well spread over the population domain. Figure 3.3 demonstrates the process following the example shown in Figure 3.2.

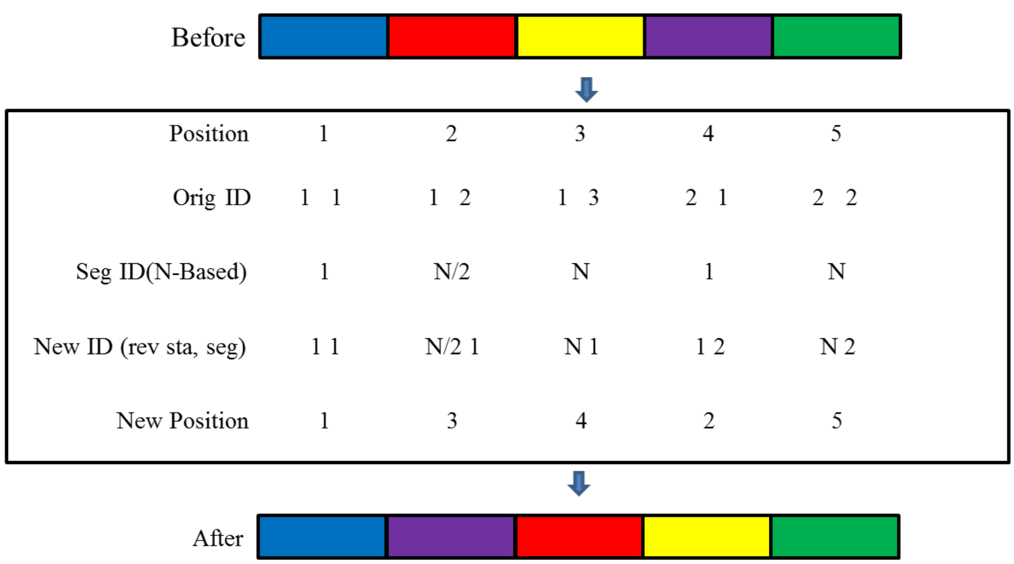


Figure 3.3 Reversed hierarchical order illustration following the example shown in Figure 3.2

4. DATA COLLECTION

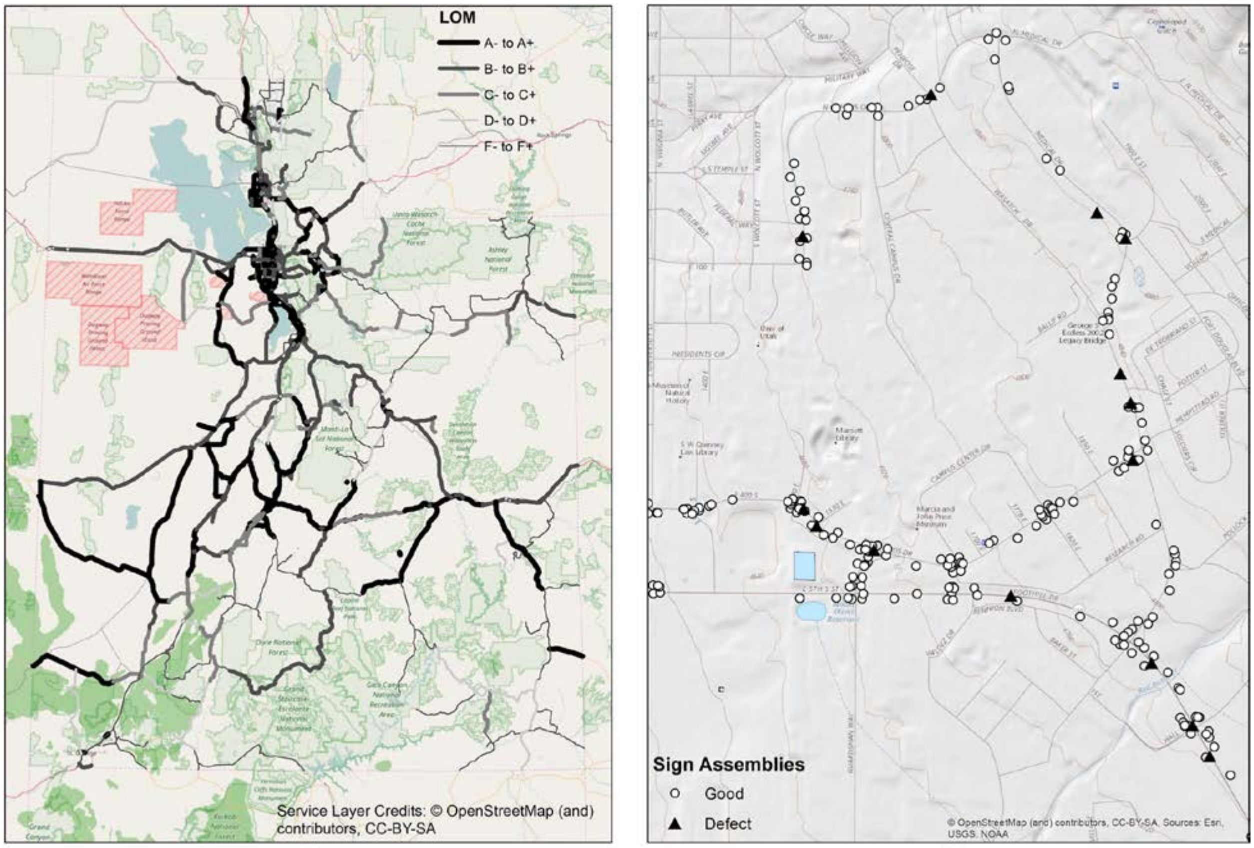
The proposed sampling methods are tested with highway asset inspection records provided by the Utah MMQA Program. Previously, MMQA performed full inventory inspection for infrastructure maintenance. The maintenance personnel recorded total numbers of infrastructures to be maintained and deficient infrastructures on each segment. Then inspection records were entered into the MMQA+ software to calculate the LOM (letter grade). One motivation for developing an asset sampling method is to reduce costs of asset inspection by estimating the overall network LOM on a sample basis. For the State of Utah, the entire highway network is divided into 489 segments. Inspection was performed semi-annually from September 2014 to March 2016, with several segments inspected multiple times in one inspection period. The inspection record archives overall infrastructure condition, segment id, infrastructure type, inspection date, and deficiency locations. There are more than 7,000 records in the database.

In fall 2014, MMQA team launched *MMQA Mobile* (Liu and Chen, 2017), an iPad application that inspection personnel use to record the defect assets. *MMQA mobile* combines field inventory with integrated global positioning system/geographic information system (GPS/GIS) mapping for maintenance data collection. Traditional data collection methods only report a total number of defects within a segment, which might result in significant bias due to human factor that cannot be validated. The *MMQA Mobile* platform, on the other hand, by enabling geotagging and description of defects on the iPad application, provides detailed geographic information of each asset's deficiency as well as the asset's condition. It adds another layer of credibility to the data, by allowing back-end post-processing to validate this data set collected from the crew for determining the LOM. In this study, the Signage Repair and Replace database is used to showcase the proposed sampling methodology. The grading scale for signage is shown in Table 4.1.

Table 4.1 LOM Grading Scale for Signage in MMQA

Percent Defect	Grade	Percent Defect	Grade
0.00-1.71	A+	13.41-14.99	C-
1.72-3.41	A	15.00-16.69	D+
3.42-5.00	A-	16.70-18.39	D
5.01-6.70	B+	18.40-19.99	D-
6.71-8.40	B	20.00-21.69	F+
8.41-10.00	B-	21.70-23.39	F
10.01-11.70	C+	23.40-100.00	F-
11.71-13.40	C		

The signage data used in this study were collected from September 2014 to March 2015 through *MMQA Mobile* at 100% signage coverage. There are 67,259 sign assemblies statewide. More than 8,500 defect observations were recorded in the database. Figure 4.1 illustrates the maintenance network with segments coded in grayscale to represent LOM during this data collection effort. A snapshot of a sample zoom-in inspection on the signs in desired/defect conditions is also shown.



(a) (b)

Figure 4.1 Location of defects and color-coded roadway segment LOM: (a) LOM of the Utah roadway network; (b) sample snapshot of the signage inspection result.

5. RESULTS AND DISCUSSIONS

In this section, we will apply the sampling methods presented in Section 3 to the maintenance records collected by the MMQA program. The analysis shows performance evaluation of the proposed sampling methods. Compared with SRS, the proposed methods show great improvements in terms of accuracy and efficiency, demonstrating their effectiveness and potential adoptability for maintenance asset management.

Figure 5.1 shows a sensitivity analysis of α, β used in Equation (4) with regards to sample size (x axis in logarithmic scale). Although the signage sampling can be considered a Bernoulli process—that takes on a value of either 0 or 1—in maintenance management, inspection is carried out on a segment basis. Since the signage condition on a segment can be considered as desired or defect, Fisher information for the Bernoulli model is adopted to segment sampling. Correspondingly, segment condition is characterized as *desired* or *defect*. With N=2,048 (number of segments) for the state of Utah, Figure 5.1 demonstrates the optimal sample size that would yield the minimum Fisher information. It shows in the figure, for example, when sample size is equal or greater than 109, the probability of sample segments being significant at 90% level is 90%.

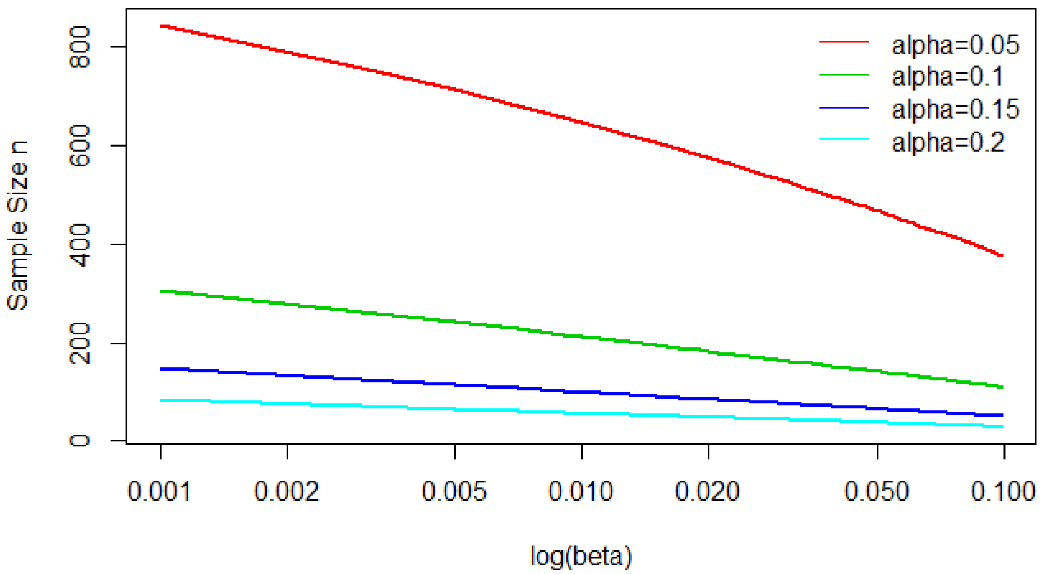


Figure 5.1 Sensitivity analysis of α, β with regard to segment sample size (x-axis in logarithmic scale).

To accommodate agencies' various sampling needs, four multi-density sampling schemes on the basis of GRTS are implemented in this study. Note that k is chosen as 500. This is selected because the minimum number of segments that a station contains is 230. This threshold ensures that the majority of stations are sampled and avoids oversampling. Four sampling methods are explored as described in the previous section, and they are *GRTS Sampling with Equal Segment Weight*, *GRTS Sampling Weighted by Signage*, *Spatially-Balanced Sampling with Equal Segment Weight*, and *Spatially-Balanced Sampling Weighted by Signage*. Note that the unequal probability is implemented to assign segment length in the one-dimensional linear structure (Figure 3.2 (c)) based on the number of signs each segment contains. Figure 5.2 presents sampling results using the foregoing four methods in the spatial context of the maintenance network. The sampled segments are highlighted in red. Note that due to the variation in segment length, the visualization might not reflect the actual sample size (with several short segments unrecognizable in the figure). The sampled number of segments for the four methods are 136, 133, 136, 134, separately, which all meet the requirements for maximum Fisher information. The latter two methods exhibit a spatially-balanced coverage.

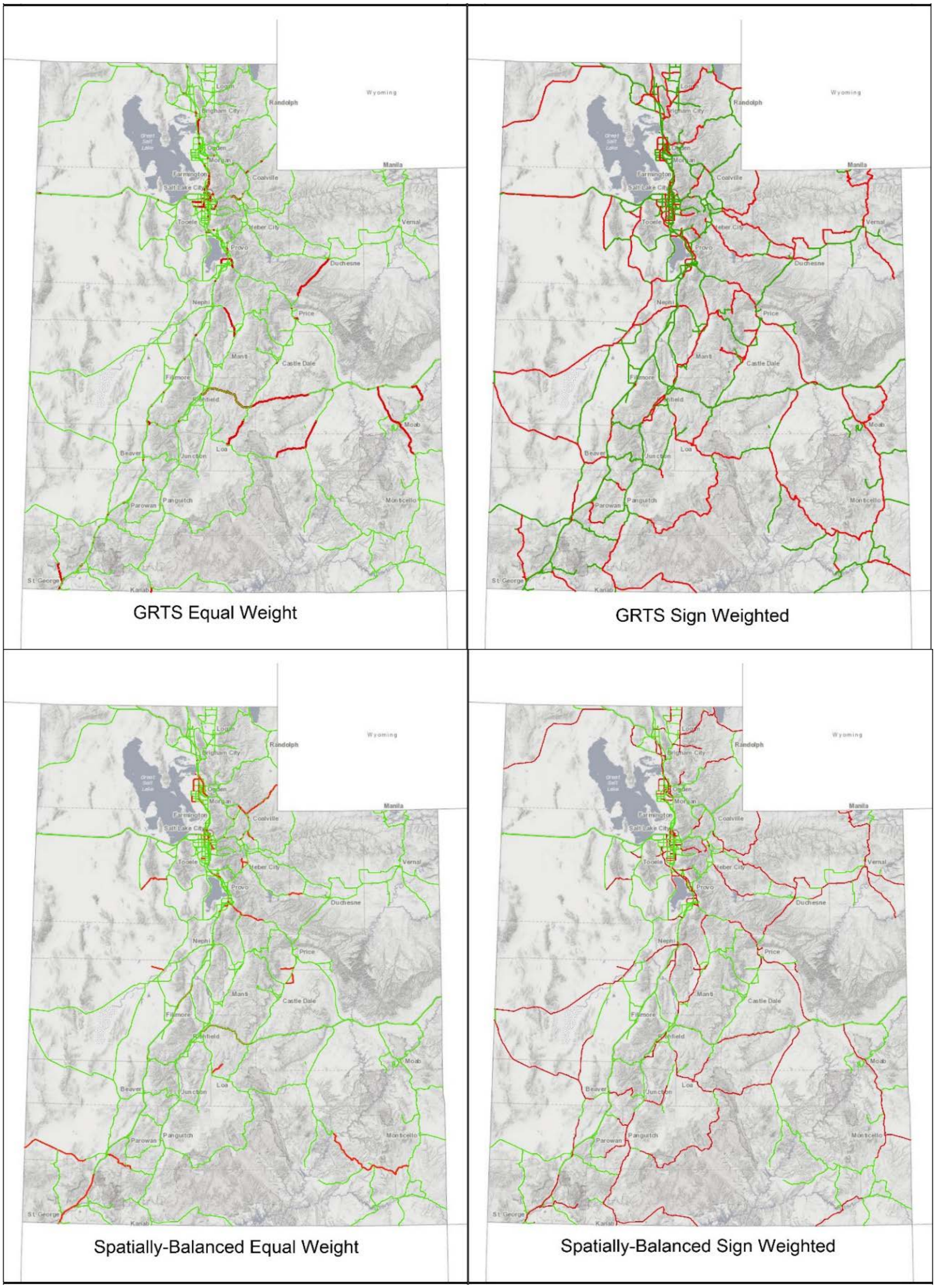


Figure 5.2 Maintenance sample segments by four sampling methods

The purpose of the sampling method development for MMQA is to provide accurate estimation of LOM at state, region, and station levels for effective budget or resource allocation. Thus, LOM can be used as an index for assessing the effectiveness of different sampling schemes when compared against ground truth statewide inventory. The *MMQA Mobile* data collected at 100% coverage is used as the ground truth data. Figure 5.3 shows the scaled LOM histogram for the sampled segments based on the four different sampling schemes compared against the ground truth data (red solid line). It is visually shown that *GRTS with Equal Weight* and *Spatially-Balanced with Equal Weight* match the statewide LOM pattern better than the rest.

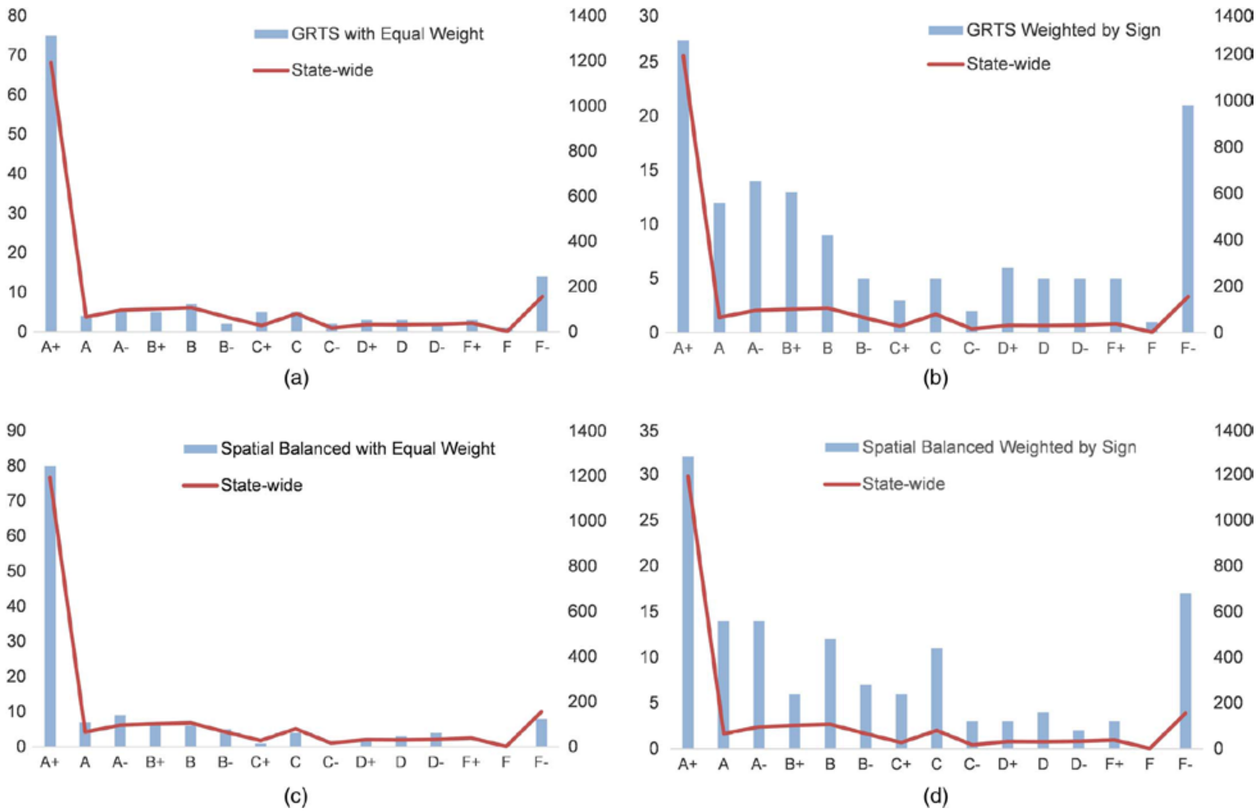


Figure 5.3 LOM histogram for the sampled segments compared against the statewide inventory.

To quantitatively measure the sampling effectiveness, a similarity analysis is conducted between statewide inventory and samples. The statewide asset LOM distribution can be represented as:

$$\mathrm{S}^{\mathrm{T}} = \left(P_{A+}^{T}, P_{A}^{T}, \dots, P_{i}^{T}, \dots, P_{F-}^{T}\right) \tag{7}$$

where $\mathrm{P}_{\mathrm{i}}^{T}$ is the true percentage of LOM i in the full inventory.

The expected LOS distribution from sample set $\mathfrak{R}$ is referred to as:

$$\mathrm{S}_{\mathfrak{R}} = \left(P_{\mathfrak{R},1}, \dots, P_{\mathfrak{R},i}, \dots, P_{\mathfrak{R},n}\right) \tag{8}$$

where $P_{\mathfrak{R},i}$ is the percentage of LOM i in sample set $\mathfrak{R}$.

The similarity measure (*d*) between sample data and ground truth data is measured using Euclidean Distance:

$$d = \sqrt{\Sigma_{i=1}^{n}\left(P_{\Re,i} - P_i^T\right)^2} \quad (9)$$

The higher *d* is, the lower the similarity between the sample set and ground truth data. We performed 30 iterative sampling runs with the four proposed methods and simple random sampling, currently widely used by transportation agencies. The similarity analysis is shown in Table 5.1. Among the four proposed methods, *Spatially-Balanced with Equal Weight* yields the best result and matches the ground truth data most closely by giving the lowest similarity score and standard deviation. Both *GRTS Sampling with Equal Segment Weight* and *Spatially-Balanced with Equal Segment Weight* methods outperform the current simple random method with much lower average similarity score. Depending on the priority or specific goals set forth by the agencies (e.g. reflect statewide LOM, station-balanced, or spatially-balanced), the appropriate sampling method can be chosen accordingly.

Table 5.1 Similarity Analysis Result: Comparing the Ground Truth Inventory with Sampling Results

Parameter	**Weighted by sign**		**Equal weighted**		
	GRTS	**Spatially- Balanced**	**GRTS**	**Spatially- Balanced**	**Simple Random**
Average	0.35588	0.35308	0.03356	0.03213	0.062627
Standard Deviation	0.02546	0.02199	0.02067	0.01493	0.017191

6. CONCLUSIONS

Maintenance management has been a critical component in today's transportation system since a sustainable network relies heavily on preservation of its infrastructure assets. An optimal maintenance program has always been focused on refining maintenance activities such that limited resources can be effectively used. The ability to report maintenance conditions with satisfying accuracy and efficiency is largely determined by data availability. Yet collecting such information can be demanding in terms of the resources, personnel, and time required. It is desirable to collect asset conditions on a sampling basis rather than on the entire asset inventory without loss of fidelity for unveiling the true asset conditions. Pertaining to maintenance management, the primary goal of sampling is to successfully estimate the overall LOM at state, region, and station levels to assist with budgeting and resource allocation. To fill this gap, this project presents a systematic approach for developing a sampling scheme customized to maintenance activities. The proposed method addresses how much and where the agencies must collect asset data with the maximum information retained for LOM estimation. The method integrates Fisher information with a spatial sampling technique that can be customized based on local agencies' requirements, such as station balanced, spatially balanced, or functional class based. These requirements are rooted in the fundamentals of maintenance management. Fisher information was applied in the study to determine asset sample size, and GRTS-based sampling methods were implemented to entertain sampling priorities set forth by the agencies. The basic idea of the GRTS method is to create a quadrant-recursive function that maps two-dimensional space into a one-dimensional one, thereby defining an ordered spatial address for the population. Unequal probability sampling can be achieved by giving each point a length proportional to its inclusion probability. Coupled with hierarchical randomization, the method is able to offer a spatially well-balanced sample.

The proposed sampling framework was showcased via an example application of the Signage Repair and Replace database maintained by the UDOT. The sampling method was enhanced on the basis of GRTS design by tailoring it to the maintenance setting. Different from the classic GRTS scheme that follows a quadrant-recursive function with the resulting address appearing as digits in a base-4 fraction, an innovative algorithm was developed to create an address for each segment with a base-N fraction given the fact that segments are already partitioned within each station with varying sizes. Four sampling methods that might be tempered to various needs were implemented, including GRTS sampling with equal segment weight, GRTS sampling weighted by signage, spatially balanced sampling with equal segment weight, and spatially balanced sampling weighted by signage. The sampling results were presented and compared against ground truth asset inventory. Comparing the simple random sampling method that is widely used by agencies across the country, both GRTS sampling with equal segment weight and spatially balanced with equal segment weight methods demonstrate better performance with much lower average similarity score. It is shown that the proposed framework lays a strong theoretical foundation for the maintenance asset sampling based on the customized requirements/needs for local agencies and is effective in estimating LOM at state, region, and station levels for budget allocation. The proposed method represents a potentially useful contribution that can be easily adoptable to any agency for optimal maintenance management. Future research includes incorporating historical LOM records into the GRTS algorithm (by weighing the sampling segments) to increase the LOM estimation accuracy.

7. REFERENCES

De la Garza, J. M., Pi˜nero, J. C., and Ozbek, M. E. (2008). "Sampling procedure for performance-based road maintenance evaluations." J. Transp. Res. Board, 2044, 11–18.

Gharaibeh, N. G., Garber, S. I., and Liu, L. (2010). "Determining optimum sample size for percent-within-limits specifications." J. Transp. Res. Board, 2151(1), 77–83.

Goodchild, M., Haining, R., andWise, S. (1992). "Integrating GIS and spatial data analysis: Problems and possibilities." Int. J. Geogr. Inf. Syst., 6(5), 407–423.

Ly, A., Marsman, M., Verhagen, J., Grasman, R., and Wagenmakers, E. J. (2017). "A tutorial on Fisher information." arXiv preprint ar-Xiv:1705.01064, Cornell Univ., New York.

Lee, M., andWagenmakers, E. (2013). Bayesian modeling for cognitive science: A practical course, Cambridge University Press, Cambridge, U.K.Liu, X., and Yu, L. (2009). "Development of geolocation data compression for transportation target identification." J. Transp. Res. Board, 2105, 71–82.

Liu, X. C., and Chen, Z. (2017). Spatial Sampling with Fisher Information for Optimal Maintenance Management and Quality Assurance. *Journal of Transportation Engineering, Part A: Systems*, *143*(10).

McCullouch, B., and Sinha, K. (2003). "Maintenance quality assurance program." FHWA/IN/JTRP-2003/18, INDOT Division of Research, West Lafayette, IN.

Medina, R. A., Haghani, A., and Harris, N. (2009). "Sampling protocol for condition assessment of selected assets." J. Transp. Eng., 10.1061/(ASCE)0733-947X(2009)135:4(183), 183–196.

Mishalani, R. G., and Gong, L. (2008). "Evaluating impact of pavement condition sampling advances on life-cycle management." J. Transp. Res. Board, 2068, 3–9.

Mishalani, R. G., and Gong, L. (2009). "Optimal infrastructure condition sampling over space and time for maintenance decision-making under uncertainty." Transp. Res. Part B: Methodology, 43(3), 311–324.

Mishalani, R. G., and Goel, P. K. (2011). Optimal condition sampling for a network of infrastructure facilities, NEXTRANS, West Lafayette, IN.

Ripley, B. (2005). Spatial statistics, Wiley, New York.

Selezneva, O., Mladenovic, G., Speir, R., Amenta, J., and Kennedy, J. (2004). "National park service road inventory program: Quality assurance sampling considerations for automated collection and processing of distress data." Transp. Res. Rec., 1889(1), 106–115.

Schmitt, R. L. R., Owusu-ababio, S., Weed, R. M. R., Nordheim, E. E. V., and Schmitt, R. L. R. (2006). Development of a guide to statistics for maintenance quality assurance programs in transportation, Midwest Regional Univ., Madison, WI

Stevens, S. (1957). “On the psychophysical law.” Psychol. Rev., 64(3), 153–181.Towsley, B., Bolot, T., Ribeiro, B., Towsley, D., Ye, T., and Bolot, J. C. (2006). “Fisher information of sampled packets: An application to flow size estimation.” Proc., 6th ACM SIGCOMM Conf. on Internet Measurement, ACM, New York, 15–26.

Stevens, D. L., and Olsen, A. R. (2004). “Spatially balanced sampling of natural resources.” J. Am. Stat. Assoc., 99(465), 262–278.

Stevens, D. L. J., and Olsen, A. (2000). “Spatially-restricted random sampling designs for design-based and model-based estimation.” Accuracy 2000: Proc., 4th Int. Symp. on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Delft University Press, Delft, Netherlands, 609–616.

Templeton, C. J., and Lytton, R. L. (1984). Estimating pavement condition and rehabilitation costs using statistical sampling techniques, NAPA, Riverdale, MD.

Myung, I. (2003). “Tutorial on maximum likelihood estimation.” J. Math. Psychol., 47(1), 90–100.

Rissanen, J. (1996). “Fisher information and stochastic complexity.” IEEE Trans. Inf. Theory, 42(1), 40–47.

Wang, J., Haining, R., and Cao, Z. (2010). “Sample surveying to estimate the mean of a heterogeneous surface: Reducing the error variance through zoning.” Int. J. Geogr. Inf. Sci., 24(4), 523–543.

Wang, J., Stein, A., Gao, B., and Ge, Y. (2012). “A review of spatial sampling.” Spatial Stat., 2(1), 1–14.

Yurek, R., et al. (2012). “Best practices in performance measurement for highway maintenance and preservation.” NCHRP Project Rep. No. 20-68A, Scan 10-03, AASHTO, Washington, DC.