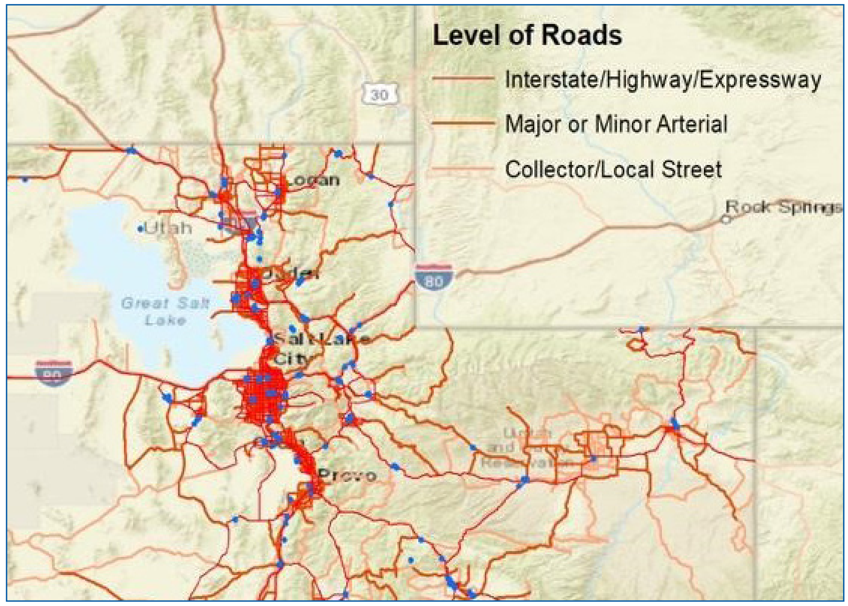
MOUNTAIN-PLAINS CONSORTIUM

MPC 21-429 | X.C. Liu and Z. Yi

LIFECYCLE ASSESSMENT USING SNOWPLOW TRUCKS' AUTOMATIC VEHICLE LOCATION DATA



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

Technical Report Documentation Page

1. Report No. MPC-544	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Lifecycle Assessment Using Snowplow Trucks' Automatic Vehicle Location Data		5. Report Date March 2021 6. Performing Organization Code
7. Author(s) X.C. Liu Z. Yi		8. Performing Organization Report No. MPC 21-429
9. Performing Organization Name and Address University of Utah Salt Lake City, UT 84112		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Mountain-Plains Consortium North Dakota State University PO Box 6050, Fargo, ND 58108		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes Supported by a grant from the US DOT, University Transportation Centers Program		

16. Abstract

Snowplow trucks serve a crucial role in winter maintenance activities by removing, loading and disposing snow. An effective performance monitoring and analysis process can assist transportation agencies in effectively managing the snowplow trucks and maintaining normal functioning of roadways. Previous literature suggests that most snowplow truck performance analysis is done through cost-benefit analysis at the macro-level to determine the optimal life cycle for the entire truck fleet. However, the proposed optimal life cycle could lead to waste of resources and may incur bias due to the ignorance of performance variations resulting from endogenous and exogenous features. More importantly, it fails to identify the contributable factors to performance deterioration. With the proliferation of data in recent years, the aforementioned concerns can be addressed through predictive machine learning techniques in a data-driven fashion. In this study, we apply machine learning techniques, including the random forest (RF) algorithm and a support vector machine (SVM) to predict the performance of snowplow trucks. Using the snowplow truck fleet managed by the Utah Department of Transportation (UDOT), both models are implemented and it is demonstrated that RF outperforms linear SVM with regard to prediction accuracy. Further, a feature importance analysis can assist transportation agencies to improve truck replacement strategy by identifying crucial factors for their performance. Lastly, a sample application of the developed prediction model suggests the threshold of work intensity for preventing rapid deterioration of trucks' performance under various working environments.

17. Key Word	18. Distribution Statement
asset management, automatic vehicle location, life cycle analysis, snowplows, winter maintenance	Public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	40	n/a

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

LIFECYCLE ASSESSMENT USING SNOWPLOW TRUCKS' AUTOMATIC VEHICLE LOCATION DATA

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

Zhiyan Yi
Ph.D. Student
Department of Civil and Environmental Engineering
University of Utah
Salt Lake City, Utah, 84112
Email: zhiyan.yi@utah.edu

March 2021

Acknowledgements

The authors acknowledge the Mountain-Plain 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:

- Jeff Casper
- Nicole Godfrey
- Vincent Liu
- Jack Mason
- Tim Ularich

Disclaimer

ABSTRACT

Snowplow trucks serve a crucial role in winter maintenance activities by removing, loading, and disposing of snow. An effective performance monitoring and analysis process can assist transportation agencies in effectively managing the snowplow trucks and maintaining normal functioning of roadways. Previous literature suggests that most snowplow truck performance analysis is done through cost-benefit analysis at the macro-level to determine the optimal life cycle for the entire truck fleet. However, the proposed optimal life cycle could lead to waste of resources, and may incur bias due to the ignorance of performance variations resulting from endogenous and exogenous features. More importantly, it fails to identify the contributable factors to performance deterioration. With the proliferation of data in recent years, the aforementioned concerns can be addressed through predictive machine learning techniques in a data-driven fashion. In this study, we apply machine learning techniques, including the random forest (RF) algorithm and a support vector machine (SVM) to predict the performance of snowplow trucks. Using the snowplow truck fleet managed by the Utah Department of Transportation (UDOT), both models are implemented, and it is demonstrated that RF outperforms linear SVM with regard to prediction accuracy. Further, a feature importance analysis can assist transportation agencies to improve truck replacement strategy by identifying crucial factors for their performance. Lastly, a sample application of the developed prediction model suggests the threshold of work intensity for preventing rapid deterioration of trucks' performance under various working environments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Snowplow trucks serve a crucial role in winter maintenance activities by removing, loading and disposing of snow. An effective performance monitoring and analysis process can assist transportation agencies in managing the snowplow trucks and maintaining normal functioning of roadways. Yet these trucks' performance could decay with the age, incurring high maintenance cost and low efficiency. It is therefore necessary to determine the optimal utilization age for the replacement of these assets. To this end, we are presenting a methodological framework using a data-driven approach to estimate the optimal utilization age of snowplow trucks, considering both total costs and operational efficiency. Specifically, a cost-benefit analysis is conducted to determine the optimal life cycle for Class 8 snowplow trucks by leveraging purchase and resale data and maintenance cost through their service span.

Meanwhile, to further analyze operational efficiency at the micro-level and to identify the crucial factors that lead to the performance deterioration, machine learning (ML) approaches, random forest (RF), and support vector machine (SVM) models, are implemented to predict truck performance using endogenous and exogenous attributes and rank the importance of those attributes. This micro-analysis can assist transportation agencies to improve truck replacement strategy by identifying key factors affecting trucks' performance. Lastly, a sample application of the developed prediction model suggests the threshold of work intensity for preventing rapid deterioration of trucks' performance under various working environments.

For this project, the Utah Department of Transportation (UDOT) provides the snowplow truck utilization data from 2000 to 2017. Exogenous features, such as weather and working environments, are collected as well for the purpose of analysis. According to the result from cost-benefit analysis, the optimal life cycle for the Class 8 snowplow trucks functioning in the State of Utah is five years. This analysis suggests a more frequent replacement cycle for snowplow trucks than what is currently implemented. Further, the annual working mileage, fuel consumption, and service year are identified as the three most important factors associated with truck performance deterioration. The results provide additional guidance on the procurement, maintenance, and replacement prioritization for Class 8 snowplow trucks.

1. INTRODUCTION

1.1 Problem Statement

Winter maintenance operations are essential to public mobility and safety, especially for areas suffering with long periods of inclement weather (Kwon and Gu 2017). In the United States, 2016 estimates suggest that annual winter road maintenance costs were $2.3 billion USD (FHWA 2016). Winter maintenance operations involve applying de-icing chemicals, snow plowing, loading snow into equipment, and hauling the snow to disposal sites (Perrier et al. 2006). To fulfill these activities, snowplow trucks are critically important. Thus, developing an effective performance monitoring and analysis process would be beneficial to the program (Adams et al. 2003). For example, a good performance from snowplow trucks ensures efficient plowing, maintains normal functioning of road networks, and avoids any potential traffic accident due to equipment malfunction. Conversely, as the equipment ages, it becomes increasingly costly to maintain the trucks due to repair costs and rapid deterioration. As a result, actively monitoring and predicting truck performance can help reduce both operation and maintenance expenses.urrently, the Utah Department of Transportation (UDOT) manages hundreds of Class 8 snowplow trucks for winter maintenance activities, including removing, loading, and disposing the snow. Generally, those snowplow trucks are sold once they are incapable of performing snowplow activities. Figure 1.1 shows the resales records of Class 8 snowplow trucks from 2000 to 2017. Notice that most trucks have a life span of over 13 years, and their life cycles are mostly concentrated between 15 to 19 years (highlighted in red in Figure 1.1). Yet with a longer service span, the truck performance is deteriorating, causing lower operational efficiency and higher maintenance costs. For instance, a portion of Class 8 snowplow trucks are equipped with “nested C-channel” frame rails. This type of frame rail can accelerate corrosion due to entrapped salts used for de-icing. Meanwhile, the repairs to frame rail cracks can be very expensive and only temporary. The maintenance costs, thus, would accumulate as trucks age, making them less reliable in servicing the roads. As a result, a method to accurately estimate the optimal life cycle to minimize the overall costs for snowplow trucks is desirable.

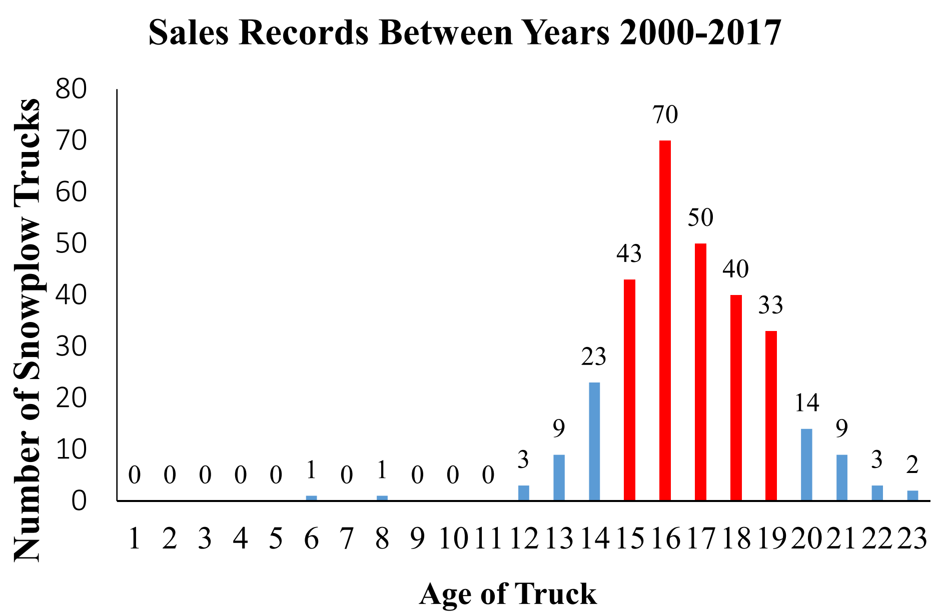


Figure 1.1 The resale records of Class 8 snowplow trucks from 2000 to 2017

Traditionally, the evaluation of snowplow fleet performance is via cost-benefit analysis, where the total cost curve across the assumed service span is constructed. This consists of equipment purchase, maintenance, and depreciation costs. The goal is to determine an optimal life cycle that minimizes overall costs while guaranteeing operational efficiency across fleet service spans (Litman 1998). Although the cost-benefit analysis helps transportation agencies identify the optimal replacement cycle for snowplow trucks, this macro-level analysis fails to evaluate the operational efficiency for a single truck, and cannot capture key factors contributing to the performance deterioration. In fact, a better understanding of truck performance can find existing trucks that may still maintain adequate performance at "optimal" replacement years determined by the model, thus avoiding the waste of resources. Moreover, it can help agencies in complementing their replacement strategy and systematically determining the service continuity/termination at the micro-level. This enables a more efficient maintenance program; one that takes advantage of variations in truck performance.

Meanwhile, other than only calculating the optimal life cycle, predicting truck performance and identifying crucial factors that lead to the performance depreciation are paramount as well. First of all, a number of trucks may still maintain decent performance at "optimal" replacement year determined by the model. As a result, replacing all trucks completely could be a significant waste of resources. Additionally, a better understanding of the performance can assist agencies in refining their replacement strategy and systematically determine the service continuity/termination at the micro-level. This would enable an efficient maintenance program that takes advantage of variations in truck performance. As a result, if truck performance can be monitored and predicted with high resolution and high accuracy, they can be replaced in time and help reduce maintenance expenses.

1.2 Objectives

This research project focuses on analyzing the life cycle and operational efficiency of Class 8 snowplow trucks in the state of Utah. Currently, UDOT is managing hundreds of Class 8 snowplow trucks. According to resale records, most of them were replaced after servicing over 13 years. However, more frequent repairs can occur, accompanied with higher maintenance costs, as truck performance decays with the service age. As a result, it is necessary to conduct data-driven analysis to evaluate the total cost and find the best replacement cycle that minimizes the overall costs of truck fleets while guaranteeing satisfactory performance during their service span.

The primary objective of this project is to develop a method to determine the optimal replacement year for the Class 8 snowplow truck fleet managed by UDOT. To achieve this, the purchase and resale information, maintenance records, and working mileage records are used in this study. Specifically, the purchase and resale records are used to estimate the cumulative depreciation cost with the increase of trucks' service span. Then the maintenance costs are utilized to reflect the maintenance expense with inflation rate at different service years. Lastly, a cost-benefit curve is constructed to decide the optimal life cycle, which minimizes overall costs for the snowplow truck fleet.

The secondary objective of this project is to inform transportation agencies in effectively managing snowplow trucks and finding suitable ML models that are capable of interpreting feature importance while having satisfactory performance prediction accuracy. To this end, this study proposes a tree-based ensemble model and a linear SVM to predict the performance of snowplow trucks, using frequency of major repairs as a performance indicator.

1.3 Outline of Report

The rest of the report is structured as follows. Chapter 2 summarizes the literature on the cost-benefit analysis for truck fleets and performance prediction analysis using RF and SVM. Chapter 3 illustrates the detailed formulation of the cost-benefit method and RF and SVM models. Chapter 4 describes the data sources used for this study, and Chapter 5 presents the results and findings. Lastly, Chapter 6 concludes the study and outlines recommendations for future research.

2. LITERATURE REVIEWS

Cost-benefit analysis enables finding the optimal replacement cycle for a fleet of assets; while RF can identify crucial factors causing the performance depreciation of the fleet. This chapter presents the summary of previous studies on snowplow truck fleet life cycle analysis and performance prediction using RF.

2.1 Snowplow Truck Fleet Life Cycle Analysis

Previous studies related to snowplow truck fleet life cycle assessment attempted to address this issue via cost-benefit analysis. With that approach, the total cost curve across the assumed service span is constructed; it consists of equipment purchase, maintenance, and depreciation costs. The goal is usually to determine an optimal life cycle that minimizes overall costs while guaranteeing operational efficiency across fleet service spans (Litman, 1998). For instance, Iowa DOT proposes a decision support system (DSS) based on a cost-benefit method to optimize the equipment life cycle (Scheibe et al. 2017). Specifically, two types of snowplow trucks are analyzed in this study, the single-axle A07 and double-axle A12 snowplow trucks, which both have a replacement cycle of 15 years. To obtain the optimal life cycles, historical data over the past nine years, including purchase price and date, maintenance cost, and actual resale values, are extracted. Additionally, they adjusted the maintenance event costs with an inflation rate of 4.23% for comparative analysis. By building total cost curves with cost-benefit method, the result suggests an optimal life cycle of eight years for A07 and six years for A12, separately. It is estimated that the Iowa DOT could save approximately $8.2 million every year by shortening the current life cycles to the recommended ones. In fact, the calculated optimal life cycle can vary greatly among different types of snowplow trucks. Wyrick and Erquicia (2008) used a similar approach to analyze the optimal life cycles for seven different types of snowplows in the state of Minnesota and conduct sensitivity analysis for purchase price, interest rate, depreciation value, and maintenance cost. The results indicate that the optimal replacement cycles range from five to 13 years based on different types of trucks. This conclusion suggests the necessity of analyzing the optimal life cycle for Class 8 snowplow trucks based on their own circumstance in the state of Utah.

Although this type of macro-level vehicle fleet analysis helps transportation agencies explore truck performance to reduce total costs, such approaches are deficient in several ways. First, existing trucks may still maintain adequate performance at "optimal" replacement years determined by the models. As a result, the complete replacement of all trucks can be a significant waste of resources. More importantly, existing cost-benefit frameworks cannot identify key factors contributing to the performance deterioration. In fact, a better understanding of truck performance can assist agencies in refining their replacement strategy and systematically determining the service continuity/termination at the *micro-level*. This enables a more efficient maintenance program; one that takes advantage of variations in truck performance. With the proliferation of data and increased popularity in data-driven approaches, it is possible to address these concerns through predictive machine learning (ML) techniques.

2.2 RF Model and Performance Prediction

ML approaches analyze data first and then automate the building process of analytical models. These models can help identify potential patterns from the data and make predictions with minimal human intervention (Alpaydin 2009). Currently, researchers use ML models in many fields, including finance, healthcare, engineering, marketing, and manufacturing, among others (Langley and Simon 1995). For example, Ghobadian et al. (2009) use artificial neural network (ANN) to predict diesel engine performance with regard to fuel consumption and exhaust emissions. Ma et al. (2017) apply a convolutional neural network (CNN) to predict traffic speed to analyze the performance of a large-scale network. Although both ANN and CNN generate high levels of predictive accuracy, they are not capable of interpreting the importance of various inputs to the results.

The Random Forest (RF) is a classic tree-based ensemble model proposed by Breiman (2001). It combines multiple base models and derives the final result via weighted or unweighted voting or averaging (Dietterich 2000). The main idea behind ensemble model is that a series of base models can generate a more stable prediction and have stronger generalization ability than a single base model. In RF, the base models are decision trees (DTs). Apart from its stable structure, another highlight of RF is that it enables interpretation of feature importance by ranking those variables based on its interior tree structure. A good illustration of this method is Bukhsh et al. (2019), where the authors use tree-based classification models, including DT, RF, and gradient boosting decision tree (GBDT), to predict the maintenance of railway switches. The results indicate that RF achieves the highest accuracy (0.70) for classifying the maintenance type among all models. In this case, accuracy is a measure of correct predictions using the model compared with the total data points. Meanwhile, Bukhsh et al. also find that functional location, service years, and detected problems are the most important features in affecting the status of switches. This interpretability can facilitate the decision-making process for infrastructure managers and help prioritize future data collection efforts.

2.3 SVM Model and Performance Prediction

Support Vector Machine (SVM) is another popular ML algorithm grounded in statistical learning theory, introduced by Cortes and Vapnik (1995). SVM takes several forms, including linear SVM and non-linear SVM with the extension of kernel trick. One advantage of linear SVM is that it reduces computational complexity as the size of the dataset increases compared with non-linear algorithms. Moreover, parameters in linear SVM are easy to interpret and help to analyze feature importance. Marković et al. (2015) measure railway performance by applying SVM to predict train arrival delays by using train category, scheduled time of arrival, headway, and other variables collected from Serbian Railways. The authors compare predictive ability to that of an ANN. The results show that the mean R^2 for SVM on the test set is 0.627, which is 4% higher than that of ANN (0.585). This result encourages further applications of SVM in performance prediction problems. Similarly, Sun et al. (2018) explore SVM's capability in predicting travelers' ticket purchase timing using high-speed rail ticket sales data in China. The authors compare results with ANN and multiple linear regression (MLR). The statistical analysis reveals that both ANN and SVM achieve the same R^2 (0.88), whereas MLR only yields an R^2 of 0.71. However, SVM produces the lowest mean absolute error (MAE) and root mean squared error (RMSE).

3. METHODOLOGY

This section describes the cost-benefit method, RF model, and SVM model, respectively. The cost-benefit method utilizes maintenance information and estimates depreciation trends of the snowplow truck fleet to pinpoint the optimal life cycle. RF and SVM models are capable of predicting truck performance by learning features related to trucks' work intensity (e.g., working mileage) and working environments (e.g., traffic volumes of the roads they serve).

3.1 Cost-Benefit Method for Snowplow Truck Life Cycle Analysis

Cost-benefit analysis takes overall costs of the assets into consideration by constructing economics curves to identify the optimal life cycle. For a snowplow truck fleet, two types of costs are mainly considered: the maintenance cost and depreciation cost. This section introduces the procedure of building the average total costs curve for a truck fleet across different service spans.

3.1.1 Maintenance Cost Calculation

Maintenance costs refer to the expenses of any required repairs as well as the costs of preventive maintenance during their service span. The maintenance costs for Class 8 snowplows are generally originated from facility replacement and labor costs with related technicians. Additionally, it is important to consider the inflation rate to adjust the costs because maintenance records span a long period of time up to 17 years. In order to construct the maintenance cost curve, the following steps are conducted:

- Aggregate the maintenance records by service years for each snowplow truck, and sum up the total maintenance costs for all trucks at each service year;
- Adjust the total annual maintenance costs with inflation rate accordingly based on different service years;
- Divide the number of trucks in service at each service year with the corresponding adjusted annual total maintenance costs.

As a result, we can obtain the adjusted average maintenance costs at different service years, which can be further used for cost-benefit analysis. It is noted that the concept of service year is different from the concept of actual year. Service year refers to the time period from the point that a truck started to service the roads. Two trucks with the same service year can be operating at different actual years. For instance, two snowplow trucks started their service at 2000 and 2001, respectively (actual year), and both functioned for one year. In this case, they have the same service year of one, but the service year happens at different actual years.

3.1.2 Depreciation Cost Estimate

Depreciation refers to the decrease in asset value in response to time. For Class 8 snowplow trucks, their values start to decrease as they start to service the road. With the increase of service span, the depreciation cost for each service year is accumulating. In other words, the assets' surplus value is decreasing with a longer service span. The relationship between the cumulative depreciation cost and surplus value can be expressed as follows:

$$p_{purchase} = sv_n + c\_dep_n \tag{1}$$

where $p_{purchase}$ stands for the original purchase price, sv_n is the surplus value at the nth service year, and c_dep_n is the cumulative depreciation costs at the nth service year.

Yet in most cases, surplus value is not available. To estimate the cumulative depreciation cost across different service years, the declining balance (DB) method is used to predict the annual depreciation value of snowplow trucks first, and then cumulative depreciation cost can be derived. DB method is an accelerated depreciation method, in which the depreciation expense is the highest in the initial year and declines over service time (Mayer 1947). The formulation for depreciation under the DB method is expressed as follows:

$$p = 1 - \sqrt[n]{\frac{s}{c}} \tag{2}$$

where p is the percentage of annual depreciation; n is the number of years of useful life; s is the surplus value at the n^{th} year; and c is the original purchase cost. Once the percentage of annual depreciation p is derived, the annual depreciation cost at a given service year (k) can be calculated based on the method as:

$$dep_k = c * (1 - p)^k \tag{3}$$

Finally, the cumulative depreciation costs from the start of service to the k^{th} service year can be derived as:

$$c\_dep_k = \sum_{i=1}^{k} dep_i \tag{4}$$

In order to obtain the percentage of depreciation p in Equation 2, we use the average resale value of trucks at a specific service year to represent the surplus value in this analysis. The detailed calculation will be presented in Section 5.

3.1.3 Cost-Benefit Analysis

Cost-benefit analysis enables the determination of optimal life cycle by observing the economic curve across different truck fleets' service spans. Once the maintenance cost curve and cumulative depreciation cost curve are constructed, the annual average total cost (AATC) per truck per mile given a specific life cycle *N* for Class 8 snowplow trucks can be formulated as follows:

$$AATC_N = \sum_{i=1}^{N} \left(\frac{Maintenance_i}{T_i} + Dep_i \right) / (N * Mi_N) \tag{5}$$

where $Maintenance_i$ is the overall adjusted maintenance costs at the i^{th} service year; T_i is the total number of trucks in service at the i^{th} service year, Mi_N is the sum of mileage records from the initial service year to the end of the life cycle. Finally, the cost-benefit curve can be plotted by calculating AATC with different life cycles *N*. The life cycle with the lowest cost in the curve is identified as the optimal life cycle (or replacement cycle) for Class 8 snowplow trucks.

3.2 RF Model

Supervised ML uses algorithms to learn the mapping function from input to output (Alpaydin 2009). In supervised learning, each sample in the dataset consists of input variables (typically a feature vector) and an output variable. The mapping function trains the algorithm so that it can use new input variables to predict the output variable with good accuracy. Classification is a subcategory of the supervised ML problem, which uses samples labeled by different classes to train mapping functions first and then predict the class of new inputs.

DT is a popular supervised learning algorithm because it is computationally efficient. The tree-shaped model can split the data by different attributes for classification or regression purposes. Generally, a decision tree consists of a root node, several interior nodes, and leaf nodes. One feeds training data into the root node and then splits these data into different interior nodes, based on data attributes. The node that helps split the dataset is called the *father node*, and the nodes after branching are called *children nodes*. If the samples in one node belong to the same class, this node terminates branching and becomes a leaf node. The tree grows recursively until all samples are assigned to leaf nodes. There are a range of additional models, built upon the basic structure of DT, that modify rules associated with data splits, branching, and subtree pruning. The most popular DT models include ID3 (Quinlan 1986), C4.5 (Quinlan 2014) and classification and regression tree (CART) (Breiman 1984) models. We detail the pseudocode for a simplified DT model in Figure 3.1.

```
define an empty decision tree T and feed the data into the root node
while True{
    select one attribute that maximizes the information gain, and
    split the data from the father node into children nodes
    if the samples are pure in the children node:
       terminate branching and form a leaf node
       if all samples are assigned to the leaf nodes:
          break
    else if all attributes are used out for splitting:
       break}
Prune the decision tree T
Return T
```

Figure 3.1 Pseudocode for a simplified decision tree model

In this pseudocode, information gain is an index to measure the classification ability given by one attribute. The larger the information gain, the stronger the classification ability. For each split, it chooses the attribute that can maximize the information gain. For instance, in CART (a binary decision tree), information gain by feature k is:

$$Information\ Gain_{(k)} = Gini_{father} - \left(\frac{|S_{left}|}{|S|} * Gini_{children(left)} + \frac{|S_{right}|}{|S|} * Gini_{children(right)}\right) \quad (6)$$

where $|S|$, $|S_{left}|$ and $|S_{right}|$ represent the number of samples in father node, the number of samples in the left children node, and the number of samples in the right children node, separately. $Gini$ is the index to describe the impurity of the node:

$$Gini = 1 - \sum_{i=1}^{N} {P_i}^2 \quad (7)$$

where P_i is the probability of the ith event happening in the node. In classification decision tree, event corresponds to the fraction of a class in one node.

However, one disadvantage of DT is the propensity to overfit the model, even when actively pruning. Moreover, the model is highly sensitive to dataset, which means that the structure of the tree may deviate significantly even when a small portion of the training data is changed. To supplement the performance of DT, one can use several tree ensemble models, including RF.

RF uses CART (Breiman 1984) as its base model. The basic idea of RF is that it generates a number of trees and combines them by weighted or unweighted averaging or voting the results from each tree. The superiority of RF is attributable to bootstrap aggregating and random feature selection techniques (Breiman 2001). The bootstrap aggregating method enables each DT to train with a subset of the data with replacement. Meanwhile, feature selection strategy allows a limited number of randomly chosen features from each tree for training. By applying these two techniques, RF is able to generate a number of different DTs and merge them into a robust tree ensemble. RF outperforms single DT by mitigating overfitting and sensitivity of the dataset effectively. Apart from that, one can generate trees in RF via parallel computing to reduce computational complexity. Figure 2 details the pseudocode for constructing a simplified RF model.

```
define a RF with M empty trees and import the dataset N
for i=Tree(1),…,Tree(M) do {
   use bootstrap aggregating method to draw samples from dataset N;
   randomly choose T' variables from T variables (T'≤T);
   train the samples with the CART model using T' variables;
 }
combine M trees together
return f(Tree(1)+…+Tree(M))
```

Figure 3.2 Pseudocode for a simplified RF model

In Figure 3.2, M is the total number of trees, N is the size of training samples, and T is the number of input variables. When one uses RF for regression tasks:

$$f\left(Tree_{(1)} + \cdots + Tree_{(M)}\right) = \frac{1}{M}\sum_{m=1}^{M} Tree_{(m)} \tag{8}$$

Alternatively, when one uses RF for classification tasks:

$$f\left(Tree_{(1)} + \cdots + Tree_{(M)}\right) = argmax_{c\in y}\left|\{m|Tree_{(m)} = c\}\right| \tag{9}$$

where y is the total number of classes in the classification task.

Besides the robust structure, another important trait of RF is its ability to interpret feature importance. In DT, every node split uses a single feature. One can compute the decrease of impurity (i.e., information gain in Equation 6) accordingly and rank features according to the average decrease in impurities across all trees in the forest. By identifying the important features in determining snowplow truck performance, transportation agencies can make informed decisions on truck replacement. For instance, if pavement conditions are the dominant force in affecting truck performance, trucks serving roads with poor pavement should operate less frequently and/or be replaced sooner.

3.3 Linear SVM

SVM is another supervised learning algorithm used for classification and regression analysis. For classification problems, given the training samples, $\{(x_1, y_1), \ldots, (x_i, y_i), \ldots, (x_n, y_n)\}$, where $x_i \in R^m$ is a feature vector with m features, $y_i \in R^1$ is the target value, n is the size of training dataset and m is the size of features, linear SVM constructs a hyperplane (or set of hyperplanes) with the largest margin to separate the samples into different classes. This hyperplane is:

$$\langle w^*, x\rangle + b^* = 0 \tag{10}$$

where $\langle w, x\rangle$ is the dot product of *w* and *x*; $w^* \in R^m$ and b^* are the parameters of the hyperplane with the largest soft margin. The corresponding classification function is:

$$f(x_i) = sign(\langle w^*, x_i\rangle + b^*) \tag{11}$$

where $f(x_i)$ is the predicting value for the i^{th} sample. One can derive parameters w^* and b^* by solving the following convex optimization problem:

$$min_{(w,b,\xi)} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \tag{12}$$

$$\text{s.t.} \qquad y_i(\langle w, x_i\rangle + b) \geq 1 - \xi_i, \quad i = 1,2,\dots,n \tag{13}$$

$$\xi_i \geq 0, \qquad\qquad i = 1,2,\dots,n \tag{14}$$

where C is the penalty hyperparameter, controlling the trade-off between soft margin and classifying training samples correctly; $\xi_i \in R^m$ are slack variables for the misclassified samples. One can transform and solve the optimization problem through its dual. The above formulation illustrates the linear classification case, where one can use the parameter w^* to interpret the importance of features, similar to a linear regression model.

In this study, both RF and linear SVM are implemented using open source package Scikit-learn in Python (Scikit-learn 2019). The number of trees M in RF and penalty hyperparameter C in SVM are determined from data with the goal of deriving an optimal classification model. We present details on parameter tuning in the *Result and Analysis* section.

4. DATA DESCRIPTION

We apply the proposed methods to examine the performance of Class 8 snowplow trucks operated and maintained by the Utah Department of Transportation (UDOT). We organize snowplow truck data and performance records, by month for each individual truck, for the years 2000 through 2017. We also extract information pertaining to downtime repair costs, working mileage, and fuel consumption per month, as well as service year and load type for each truck. The data are high quality, with approximately 0.21% erroneous recording on mileage of extremely large values. We perform additional filtering of data through interpolation to replace those outliers with averaged odometer records of previous and following months. We aggregate endogenous variables, such as working mileage and fuel consumption, into average annual consumption across each truck's entire service span to implement the RF and SVM models.

4.1 Dataset for Cost-Benefit Analysis

The AATC curve for cost-benefit method consists of the maintenance cost and depreciation cost. To derive AATCs with the variation of life cycles, a series of utilization information for Class 8 snowplow trucks is needed. The detailed descriptions of used datasets are listed below:

- **Maintenance costs records**
  The maintenance costs for each snowplow truck are recorded monthly from 2000 to 2017. Maintenance can be classified as commercial repairs and non-commercial repairs, where commercial repairs denote restorations by professional technicians from third party companies and non-commercial repairs are restorations by UDOT. The maintenance cost is aggregated by service year to obtain the optimal life cycle.

- **Number of trucks**
  The number of trucks serving at the same service year is not equivalent to the number of trucks serving at the same actual year. This information is required for calculating the annual average total cost per truck. The records indicate that there are 831 Class 8 snowplow trucks in total serving the roads from 2000 to 2017 in Utah. The oldest snowplow truck started its service in 1979 and the latest one started in 2017. For the first service year, there are 473 trucks in total. However, the number of trucks become fewer with the longer service span. At the 17th service year, there are only 32 trucks with available service records.

- **Mileage records**
  The mileage records for each snowplow truck are documented monthly. We sum up the mileage data by each service year to calculate AATC per truck per mile. The provided data are determined to be of high quality with approximately 0.21% erroneous recording on mileage of extremely large values. Further filtering is conducted through interpolation to replace those outliers with averaged odometer records of previous and following months.

- **Original purchase data**
  The purchase data include original purchase price and purchase date for each truck. There are 521 Class 8 snowplow trucks being purchased through the years 2000 to 2017 in total. The purchase records are used for DB method and inflation rate evaluation to adjust maintenance costs in different actual years.

- **Disposition information**
  The Class 8 snowplow trucks will be sold once they cannot function normally and serve the roads. Resale information contains resale price and date. The disposition data are utilized to estimate the percentage of annual depreciation for DB method. According to the dataset, there are 301 Class 8 snowplows being sold with valid information from 2000 to 2017.

- **Label of working regions**
  The Class 8 snowplow trucks are assigned to different working regions in Utah. Each truck is labeled with a specific working region. The four working regions are Salt Lake City, Ogden, Orem, and Richfield. Cost-benefit analysis can be performed based on different regions to explore the variation of optimal life cycles across geographical areas.

4.2 Dataset for Truck Performance Prediction

For prediction purpose, information such as repair costs for each downtime, working mileage and fuel consumption per month, service year and load type for each truck is extracted. Endogenous variables such as working mileage and fuel consumption are further aggregated into annual average consumption across each truck's entire service span to implement the proposed method. Other than the endogenous variables that could impact snowplow truck performance, we hypothesize that exogenous factors such as weather and terrain type could also potentially affect their operation. We collect a range of additional data for modeling, including snow depth records from 2000 to 2017, functional classifications and annual average daily traffic (AADT) of roads in Utah, and land use data indicating whether the truck serves urban or rural areas. Meanwhile, all snowplow trucks are equipped with Verizon Automatic Vehicle Location (AVL) technology. Verizon AVL is a fleet tracking system that enables monitoring and managing mobile equipment. This GPS fleet tracking system records truck locations and speed information every two minutes, which allows near real-time monitoring of the fleet. To delineate the territorial and/or land-use impact on snowplow truck performance, we retrieved the origin and destination (OD) data of each active trip for every truck between February 6 and March 31, 2018 – the period when trucks were performing major winter maintenance activities. An active trip refers to the daily trajectory of snowplowing activity for each truck. Table 4.1 lists detailed information of all variables collected.

Table 4.1 Detailed description of all variables

Variable	Denotation	Description	Unit	Resource
Year	service year	The service year is determined from the start year to 2017 if the truck is still in service, otherwise it is determined from the start year to the year that it is sold.	year	UDOT
Fuel	average annual fuel consumption	It records the average annual fuel consumption for each truck during its service span.	gallon/year	UDOT
Mi_winter	average annual mileage in winter season	The average annual mileage for each truck in winter season during its service span. Winter season is defined from November to March (in the year to follow).	mile/year	UDOT
Mi_other	average annual mileage in other seasons	The average annual mileage for each truck in other seasons during its service span. This is to delineate other operational-similar activities that a snowplow truck might take on outside the winter season (e.g., the fire truck).	mile/year	UDOT
Type	load type	The Class 8 snowplow trucks are classified into three types, Type 104, Type 113, and Type 168 with numbers representing the capacity for snow.	NA	UDOT
Func	functional classification of the roads	The roads are categorized into seven levels, with 1 representing the highest level and 7 representing the lowest level based on the functional classification: 1: Interstate 2: Other freeway and expressway 3: Principal arterial 4: Minor arterial 5: Major collector 6: Minor collector 7: Local streets	NA	UDOT
Vol	AADT in 2016	The AADT in 2016 is used to reflect the volume of each road.	veh/day	UDOT
Snow	average annual snow depth during truck's service span	We average the snow depth for each truck during its service span to reflect the workload across different trucks.	ml	MesoWest
Area	service area type	The snowplowing activity region is distinguished by urban vs. rural regions.	NA	CTPP
Rank	rank of major repair times	The rank of major repair times is used to quantify the performance of snowplow trucks. It is categorized as follows: Rank 1: 0-4 times Rank 2: 5-8 times Rank 3: 9-12 times Rank 4: over 12 times	NA	UDOT

After acquiring exogenous features, we associate these data with each truck. For example, we link exogenous features with proximal trucks using AVL OD data to estimate the centroid of snowplowing activity. As mentioned earlier, the OD data include the OD points for daily active trips of trucks spanning February 6 to March 31, 2018. Note that most daily snowplowing activity is a round trip, where the OD coordinates are quite close for each truck. Thus, we average the coordinates of all OD points for each snowplow truck to approximate the centroid of snowplowing activity. Once the centroid is determined, we use different ring buffers with varying radii (i.e., 2km, 5km, and 10km) to capture the roads that each truck serves (Figure 4.1 a, b). The results suggest that for 95% of trucks, a 2 km buffer is sufficient to capture roads serviced. We label road functional classifications from 1 to 7, with 1 representing the highest level of mobility and 7 representing the lowest level of mobility. We average the functional classification variable for the roads within the buffer to better represent their characteristics. We replace the average functional classification values for the remaining 5% of trucks with the average values of the other 95%.

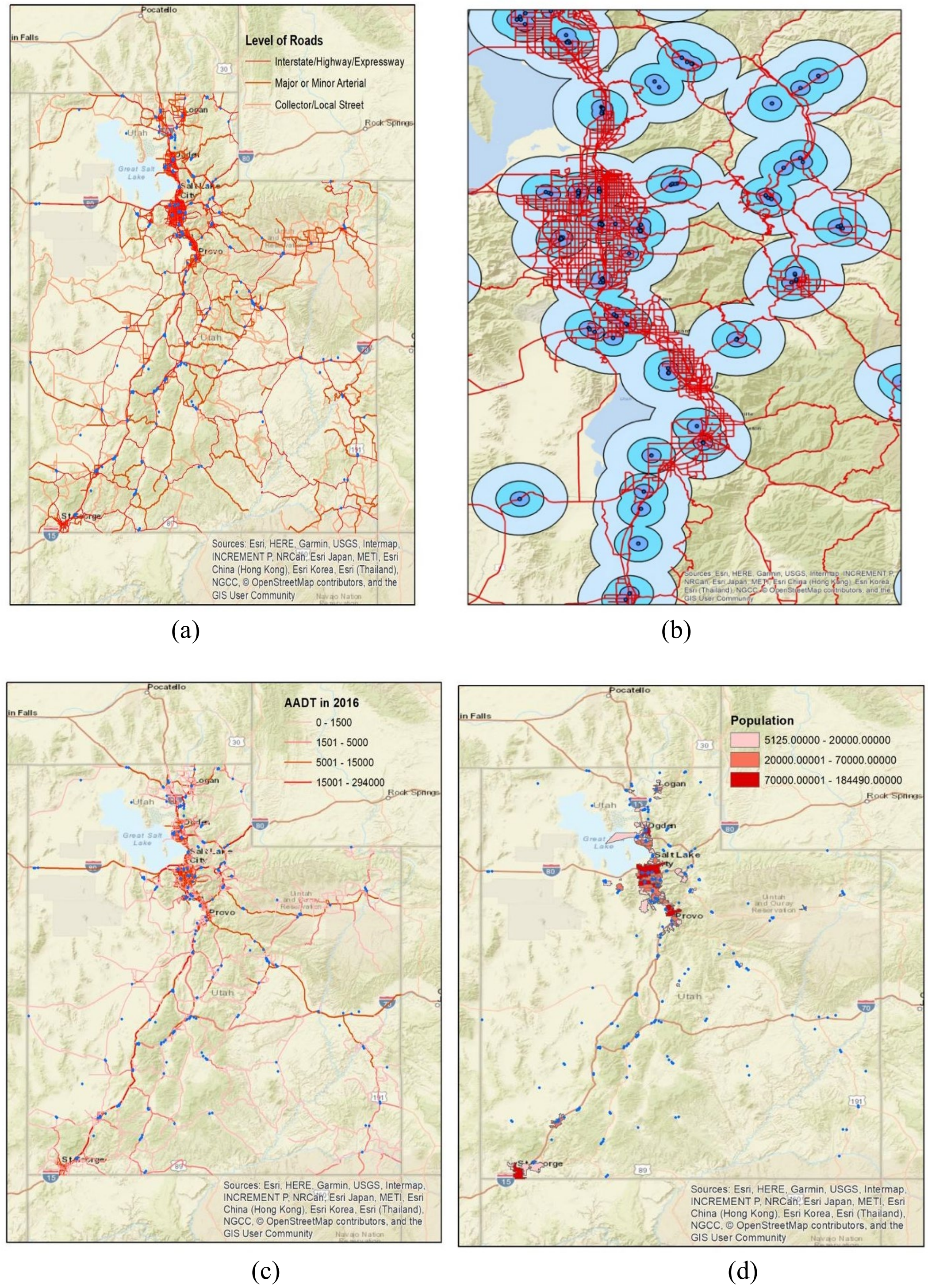


Figure 4.1 (a) Functional classification of the roads and the distribution of the centroids of snowplow activities; (b) a sample of snowplow trucks with ring buffers; (c) 2016 AADT of all roads in the state of Utah; and (d) distribution of urban area in the state of Utah.

We extract the 2016 AADT of roads that each truck serves using a similar approach as described above. The land use types are roughly classified into rural vs. urban, with urban area defined as a census tract that has a population of more than 5,000 people – those not meeting that criteria are classified as rural areas (Hall et al. 2006). We use centroid activity locations to determine land use features. We also extract population data in the state of Utah from 2006 to 2010 from the Census Transportation Planning Products Program (CTPP), a state DOT-funded cooperative program (CTPP 2019). Figure 4.1c,d display the AADT and urban area designations. Lastly, we retrieve snow depth for each year from MesoWest, a program started at the University of Utah providing access to the current and historically archived weather observations (MesoWest 2019). We average snow depth across the service years for each truck.

In this study, we use the frequency of major repairs as the main indicator to classify the performance of all trucks. For the purposes of this study, this makes sense because we are seeking to maximize efficiency and minimize repair times (Swanson 2001). Previous studies show that a single major repair for snowplows (e.g., replacement of the axles) is around $2,500 on average (Cuelho and Kack 2002). For simplicity, we define major repairs as any repair record that exceeds $2,000. Based on the range of frequency of major repairs for all trucks, we categorized them into four ranks, with detailed classification shown in Table 4.1. Overall, Rank 1 represents trucks with good performance, while Rank 4 indicates poor performance.

There are 829 trucks in service from 2000 to 2017. A portion of the trucks (420) started their service prior to 2000. Due to data unavailability, we remove these trucks from the modeling effort. The OD data from AVL missed activity records of 21 trucks due to possible causes such as resales and temporary maintenances. Ultimately, we use 388 snowplow trucks with complete records for this study.

4.3 Data Post-processing

Data post-processing involves data cleaning, normalization, transformation, feature extraction, and the like. Post-processing can have a significant impact on generalization performance of supervised ML algorithms (Kotsiantis et al. 2006). The following subsections present the data post-processing we conducted for RF and linear SVM.

4.3.1 Imbalance of Samples

In order to predict truck performance, we categorize trucks into four groups based on the severity of their repairs. For the entire fleet, the numbers of trucks in Ranks 1 to 4 are 264, 83, 34, and 4, respectively. The imbalanced classes could result in the model downplaying features in the minority classes. To fix this issue, we use resampling technique to sample the minor classes with replacement until the sample sizes are approximately equal across classes (Van Hulse et al. 2007). Consequently, the final dataset includes information of 997 trucks. The number of trucks in Rank 1 remains the same, while the numbers of trucks in Ranks 2 through 4 are populated as 249, 244, and 240, separately.

4.3.2 Normalization

Since SVM is sensitive to the scalar of features, we normalize the variables as:

$$x'_d = \frac{x_d - \mu_d}{\sigma_d} \tag{15}$$

where d represents the i^{th} variable, μ is the mean value and σ is the standard deviation for the i^{th} variable. The normalization process is only required for SVM but not RF, as RF is not sensitive to feature scaling.

5. RESULTS AND ANALYSIS

5.1 Cost-Benefit Analysis for Optimal Life Cycle

5.1.1 Inflation Rate and Maintenance Cost

Inflation is a sustained increase in the general price level of goods and services in an economy over a period of time. According to the resale records, the average purchase price for one Class 8 snowplow truck is below $100,000 in 2000, while the average purchase price reached approximately $150,000 per truck in 2017. This change indicates the necessity of taking inflation into account when calculating the total costs throughout its entire life span. For maintenance cost records, they can be at the same service year but at different actual years. As a result, those maintenance costs need to be multiplied by the inflation rate of the corresponding years. Iowa DOT's study (Scheibe 2017) used an annual inflation rate of 4.23%. To derive the fitted inflation rate in this study, we use the average purchase price in 2000 as the base, and predict the purchase price for the following years with various inflation rates (i.e., 4%, 5%, and 6%) up to 2017. The results are presented in Figure 5.1. In this figure, it can be observed that if setting 4% as the inflation rate, the purchase prices are underestimated for most of the years; while 6% of inflation rate overestimates the purchase price significantly as the service cycle increases. Therefore, 5% is considered a reasonable proximate.

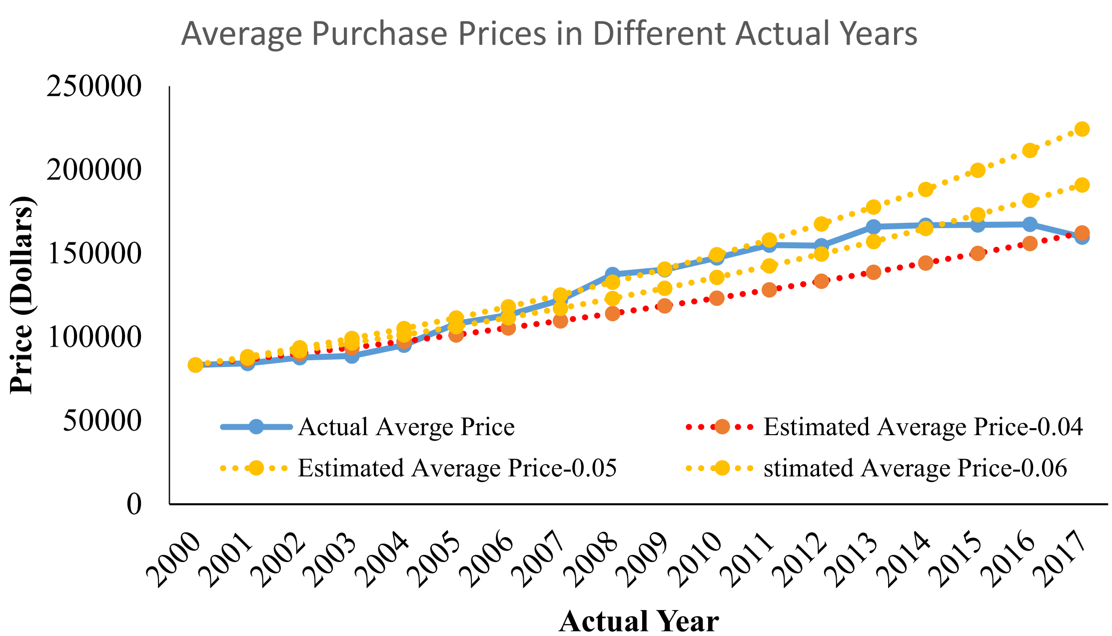


Figure 5.1 The average purchase values and estimated purchase values with different inflation rates

In the following step, maintenance costs in different years are adjusted accordingly based on the records of their actual years. In this study, year 2017 is set as the base, and all maintenance costs recorded from the previous years are adjusted with the formulation below:

$$Maintenance_{adjusted} = Maintenance_i * (1 + 0.05)^{2017-i} \tag{16}$$

where $Maintenance_i$ represents the actual maintenance costs recorded at the i^th actual year. The aggregated maintenance costs in all service years (from one to 16 years), the average maintenance costs per truck per mile in different years, and the cumulative average maintenance costs are shown in Table 5.1.

Table 5.1 The detailed information regarding maintenance costs

Service year	Number of trucks	Total cost with inflation	Total mileage	Average cost (per unit per mile)	Cumulative average Cost
1	451.00	2141821.75	4720914.00	0.45	0.45
2	457.00	2120095.47	4692610.00	0.45	0.91
3	465.00	2782768.09	4765585.00	0.58	1.49
4	480.00	3376872.04	4595855.00	0.73	2.22
5	462.00	3603744.39	4498135.00	0.80	3.03
6	462.00	4048765.37	4104520.00	0.99	4.01
7	471.00	4324249.44	4023063.00	1.07	5.09
8	476.00	4425360.46	3683168.00	1.20	6.29
9	436.00	4211407.75	3307228.00	1.27	7.56
10	436.00	4293625.37	3167493.00	1.36	8.92
11	416.00	4190546.59	2869841.00	1.46	10.38
12	405.00	4059009.48	2510458.00	1.62	11.99
13	388.00	3770801.03	2299985.00	1.64	13.63
14	379.00	3520676.26	1938350.00	1.82	15.45
15	365.00	2787833.64	1620040.00	1.72	17.17
16	304.00	2044549.29	1051591.00	1.94	19.11

5.1.2 Depreciation Cost and Surplus Value Curves

To obtain a depreciation curve, the percentage of depreciation in Equation 2 needs to be determined first. In this study, we decide to use the average resale values at the 16th service year, which is $8,109, as the surplus value *s*, and the years of useful life *n* is 16, correspondingly. This is because the 16th service year has the highest number of resale records, as shown in Figure 1.1, which provides a larger sample size on the trucks' surplus value. Meanwhile, the average purchase price in 2017 ($159,750) is used to indicate the actual purchase value *c* in Equation 2, since all maintenance costs are adjusted to the costs in 2017. As a result, the calculated percentage of depreciation *p* is 17%. Once *p* is calculated, the depreciation cost in each service year and the cumulative depreciation cost can be calculated. The snowplow trucks' cumulative depreciation costs and surplus value with the increase of service span are illustrated in Figure 5.2.

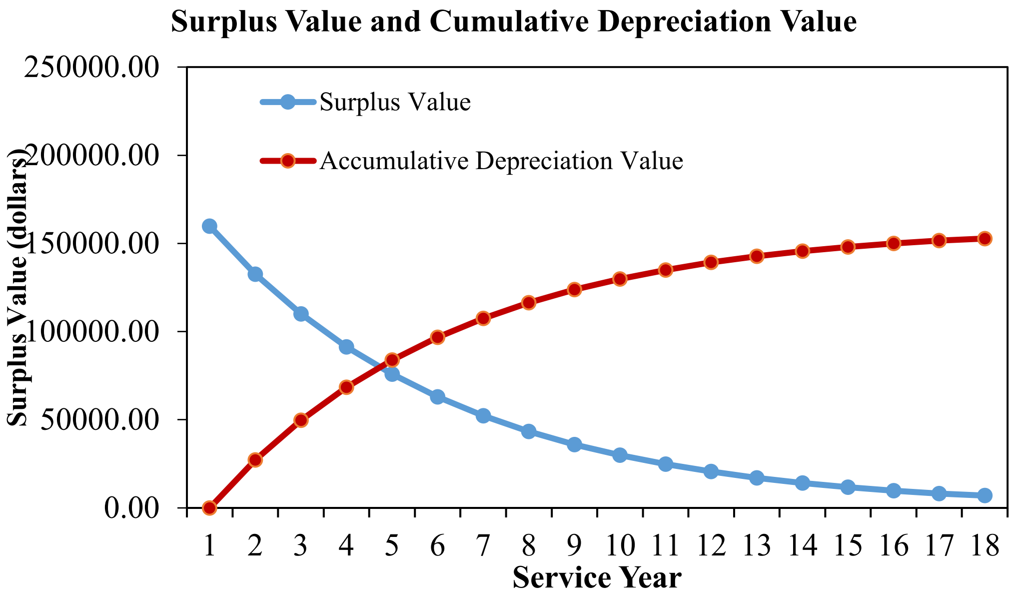


Figure 5.2 The accumulative depreciation costs and surplus value curves

Figure 5.2 shows that the trucks' surplus value continuously drops as service span increases. Since the DB method follows an exponential decay in value, the decrease of surplus value in the first several years is very pronounced. At the seventh service year, a snowplow truck remains merely one-third of its total value. Yet the surplus value steadily declines after that due to entity's low remaining values. This indicates that if the snowplow truck fleet is replaced frequently, the total cost can be extremely high because of the initial exponential decay in surplus values. Hence, the optimal total cost should balance between the depreciation cost and maintenance cost.

5.1.3 Total Cost Curve and Optimal Life Cycle

In the last two subsections, maintenance cost in different service years and cumulative depreciation cost are obtained. In this subsection, Equation 5 is implemented to calculate AATCs per truck per mile with different service spans. AATCs for life cycle from one year to 16 years are calculated and presented in Figure 5.3.

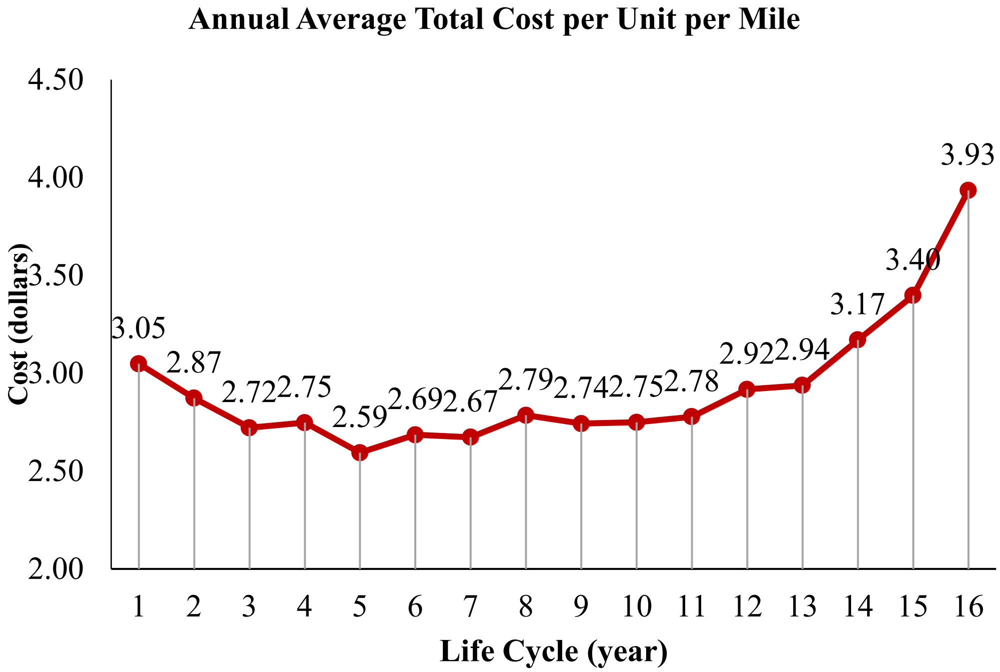


Figure 5.3 AATC per truck per mile with different life cycles

Figure 5.3 indicates that the AATC drops first and then increases gradually with the increase of life cycle. For the initial service years, although annual depreciation cost is relatively high, maintenance cost is also marginal. This is because assets usually maintain satisfactory performance as they first start the service. However, with longer service years, the maintenance costs start to mount up due to more frequent malfunctions and repairs. The optimal life cycle can be easily identified as five years, where the AATC is \$2.59 per truck per mile. Yet the current service ages for Class 8 snowplow trucks are mostly concentrated between 15 to 19 years. Assume there are 500 snowplow trucks operating annually with average working mileage of 8,000 miles per truck, if the life cycle for all trucks has shortened from 15 (AATC is \$3.40 per truck per mile) years to five years, UDOT can save approximately \$3.24 million every year. The implementation of recommended replacement strategy thus could result in significant cost savings.

In fact, the snowplow trucks are all assigned with specific working regions. Due to the terrain difference and miscellaneous reasons, the optimal life cycles may vary across different geographical areas. As a result, we perform the cost-benefit analysis based on different regions in Utah. The trucks' region classification information is extracted from the Verizon AVL trajectory database from January to March 2018. Note that the data do not include all the trucks used for the statewide analysis due to the relatively short period of time; only 471 snowplow trucks are recorded during those two months, yet 831 trucks' records are used for the statewide analysis from 2000 to 2017. However, the results can still provide much valuable information on how the performance might differentiate across regions. Figure 5.4 shows the number of snowplow trucks in each region, namely Salt Lake City, Ogden, Orem, and Richfield regions, and Figure 5.5 illustrates their corresponding activity distribution.

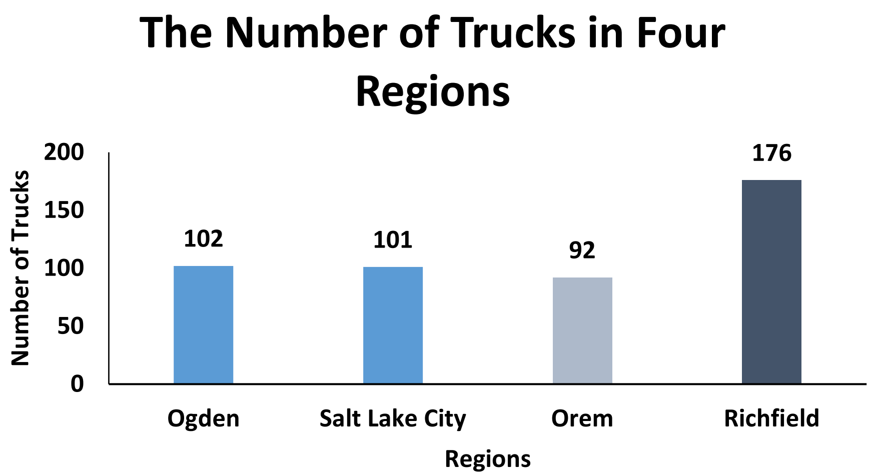


Figure 5.4 The number of snowplow trucks in four regions

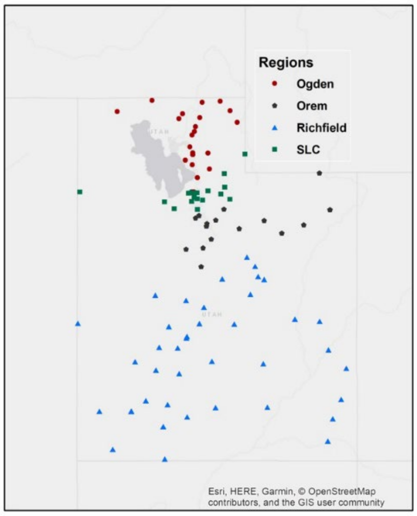


Figure 5.5 Distribution of snowplow trucks in four regions

Figures 5.6 through 5.9 further delineate AATC values by the four different regions.

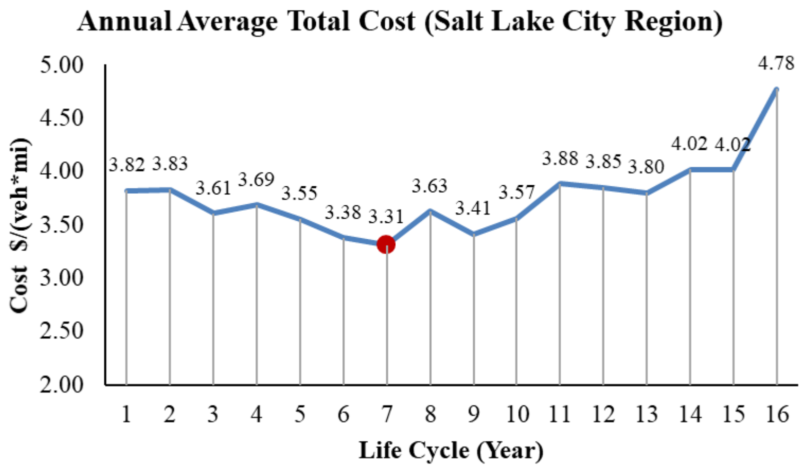


Figure 5.6 AATC per truck per mile in different replacement years (Salt Lake City Region)

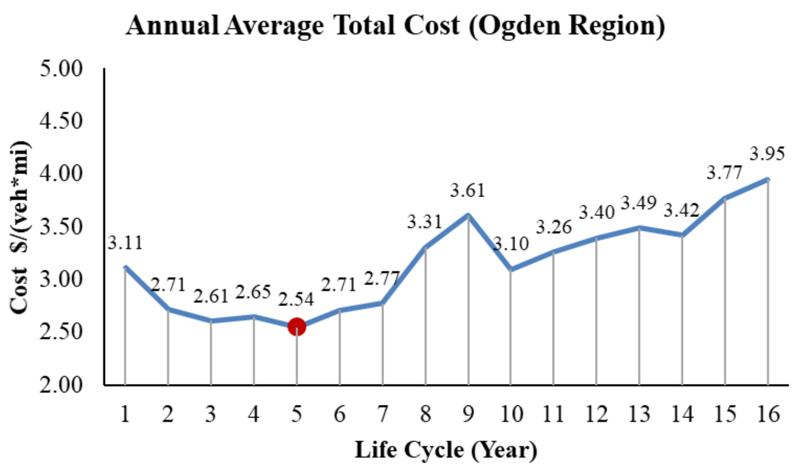


Figure 5.7 AATC per truck per mile in different replacement years (Ogden Region)

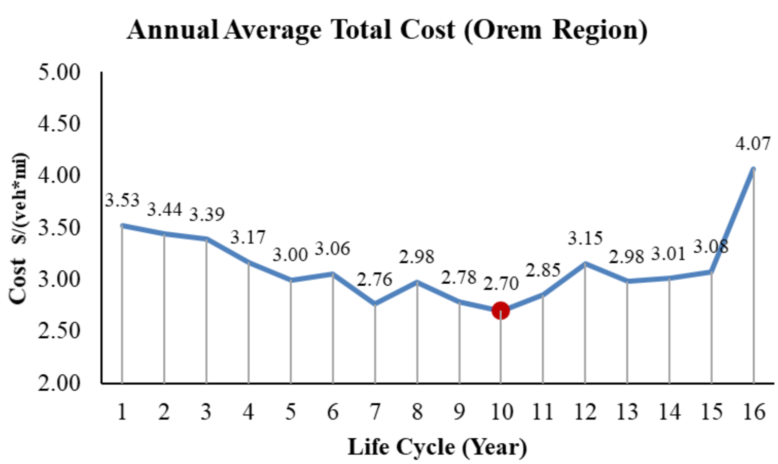


Figure 5.8 AATC per truck per mile in different replacement years (Orem Region)

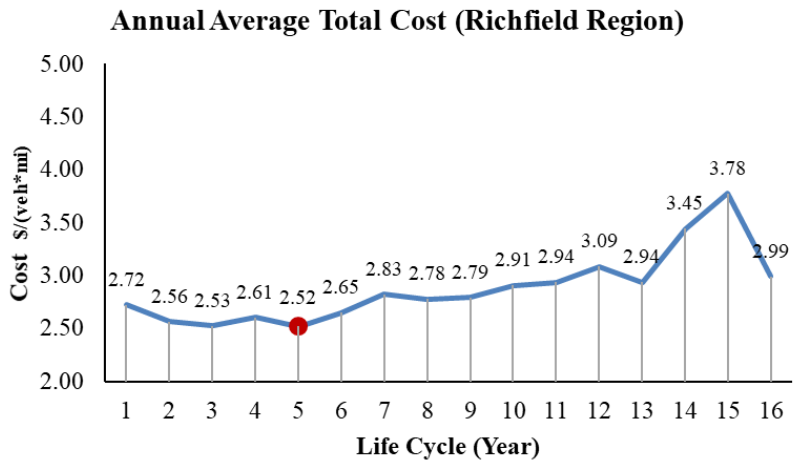


Figure 5.9 AATC per truck per mile in different replacement years (Richfield Region)

Figure 5.10 presents the previous results on the same horizon, and Figure 5.11 illustrates the comparison using the statewide value as a benchmark.

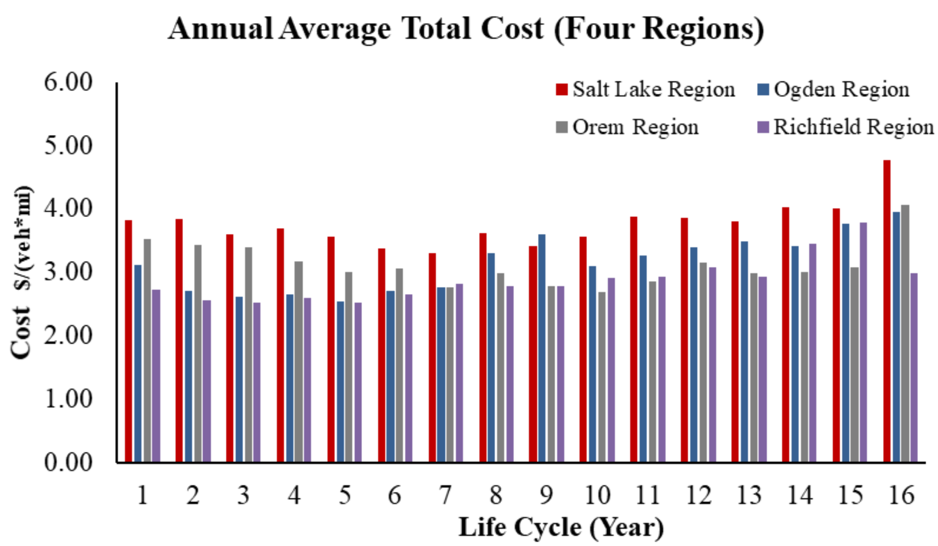


Figure 5.10 AATC per truck per mile in different replacement years (four regions)

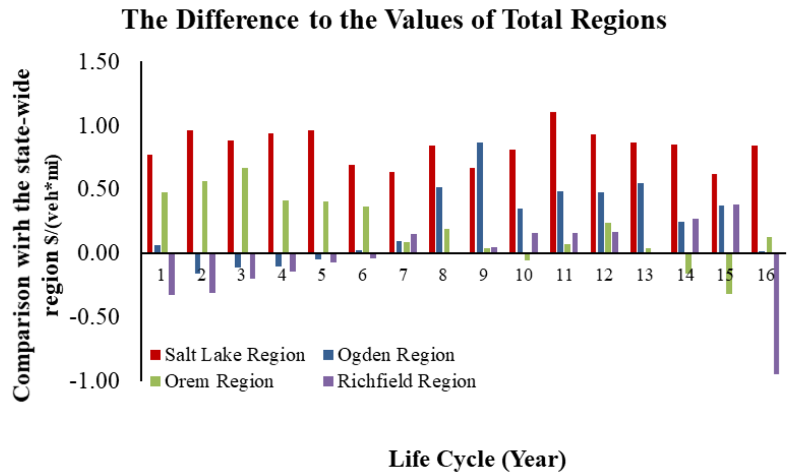


Figure 5.11 Cost comparison using statewide value as a benchmark (four regions)

Based on the results presented above, there are several findings worth mentioning. First, the optimal replacement years for the State of Utah is determined as five years. However, when we split the analysis into four regions, the results vary. Ogden and Richfield regions are still consistent with the state-wide result (five-year as optimal replacement cycle). The Salt Lake City region has an optimal year at seven and Orem is postponed to 10. This result indicates that the snowplow trucks which serve the Salt Lake City or Orem regions can be potentially extended for their lifetime service longer. Note in Figure 5.11 that the trucks serving the Salt Lake City region have the highest cost across all service years, and the values surpass the statewide average. This suggests that those trucks are relatively costly to maintain. On the contrary, trucks serving the Richfield region have relatively lower cost.

5.2 Prediction Results for RF and SVM

5.2.1 Parameter Tuning and Performance Measurement

Standard practice of ML involves splitting the dataset into a training set and test set, where one uses the training set to train the model and test set to evaluate model performance using an unknown dataset. Before training, some hyperparameters (e.g., number of trees in RF and C in SVM) are chosen manually to achieve better predictive performance. However, we run the risk of overfitting if we tune the hyperparameters on the test set. Specifically, information on the test set can "leak" into the model, reducing capabilities for generalization. To address this issue, another portion of the entire dataset, usually referred to as a "validation set," is held out to enable the evaluation of trained models and to choose the optimal hyperparameters. We then can apply the models with the best hyperparameters to the test set to report on generalization performance. One potential drawback of this method is that it reduces the size of training data by partitioning the available data into three parts (i.e., training, validation, and test sets). One remedy for this is to use the K-fold Cross-Validation (CV) (Kohavi 1995). With K-fold CV approach, the data are still split into training set and test set, but we further partition the training set into K subsets with equal size. Each time, we use one subset as a validation set, and we use the remaining K-1 subsets to train the model. We repeat this process K times and select the hyperparameters with the best average performance on validation set. Finally, we apply the selected hyperparameters to the test data. Empirically, K is typically set as 5 or 10 since it can lead to less bias and reduce the computational cost (Rodriguez el al. 2009). In this paper, we shuffle the data and split them into 80% as training data and 20% as test data, and K is set as 5.

Two performance measurements (i.e., classification accuracy score and confusion matrix) are used for this multi-classification prediction. Generally, one solves a multi-classification problem by transforming it into a binary classification problem through *one vs. all* strategy. For a binary classification problem, one can divide all samples into two classes (one class is identified as positive class and the other one as negative class). *One vs. all* strategy assumes one class as positive class and other classes as negative class. We will have N binary-classifiers, where N is the number of classes. In this study, N=4 as all samples are classified into four classes (Ranks 1 through 4). The classifier will output a list with each number in the list representing prediction probability to the index-based class. The final predicted class of the input is the class with the highest prediction probability. For example, when predicting the i^{th} sample's repair rank, if the prediction output is [0.3, 0.4, 0.8, 0.2] for Ranks 1 through 4, separately, the repair rank will be predicted as Rank 3. To measure the performance of the model, we use classification accuracy scores to indicate the models' predictive capability:

$$S = \sum_{k=1}^{K} \frac{N_{k\_correct}}{N_k} \tag{17}$$

where K is the number of classes, $N_{k\_correct}$ is the total number of samples correctly predicted as class k, and N_k is the total number of samples that are actually in class k. The classification accuracy score measures the correct predictions of the model compared to the total number of data points (Buksh et al. 2019).

Before we compare the performance of the two models, we optimize hyperparameters for each model to yield the lowest prediction error. For RF, the number of trees M is manually adjusted from 1 to 1,000. Figure 5.12a highlights the training and validation curves. Hyperparameters with the highest average accuracy score on the validation set across the 5-fold CV are selected for the final models. The validation curve shows that when the number of trees M is above 50, the average classification accuracy scores on validation set are higher than 0.9, and the classification ability does not improve much beyond 500. We therefore set M as 500 to avoid overfitting.

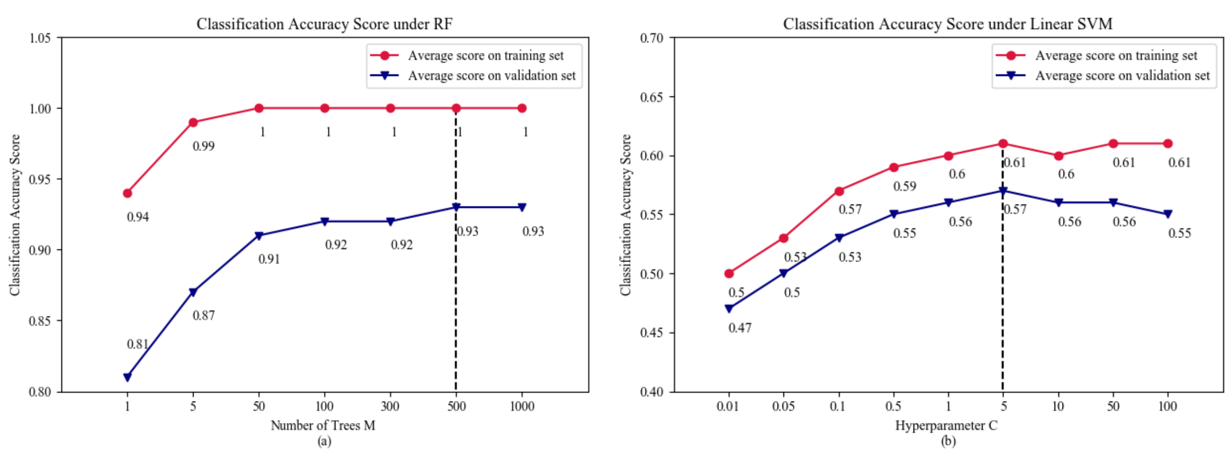


Figure 5.12 Average classification accuracy scores under (a) RF and (b) linear SVM by performing 5-fold CV

For linear SVM, we only need to tune the penalty hyperparameter C in Equation 7 ($C\epsilon[0.01, 100]$). Hyperparameter C regulates the soft margin of different classes by controlling the number of support vectors. If C is set too small, it will lead to underfitting. On the contrary, large C can cause weak generalization ability to the unknown data. Figure 5.12b shows the accuracy curves of linear SVM using 5-fold CV. The best performance on the validation curve is achieved when C is 5. Nevertheless, both test and validation curves suggest linear SVM performs relatively poorly with the highest score on the validation set reaching only 0.57. After we select parameters for both models, we execute the final models on the test set. The accuracy scores of RF and linear SVM on the test set are 0.92 and 0.51, separately. To measure the performance of the classification model more explicitly, we use a confusion matrix on the test dataset to illustrate the finer details of results using different methods. Confusion matrices allow for the identification of the success (or failure) of classification results. Specifically, each row represents the actual class, while each column represents a predicted class. Figure 5.13 illustrates the confusion matrices for this multiclassification problem. As an example, the number in the first row and second column corresponds to the instances that actually belong to Rank 1 but are misclassified as Rank 2.

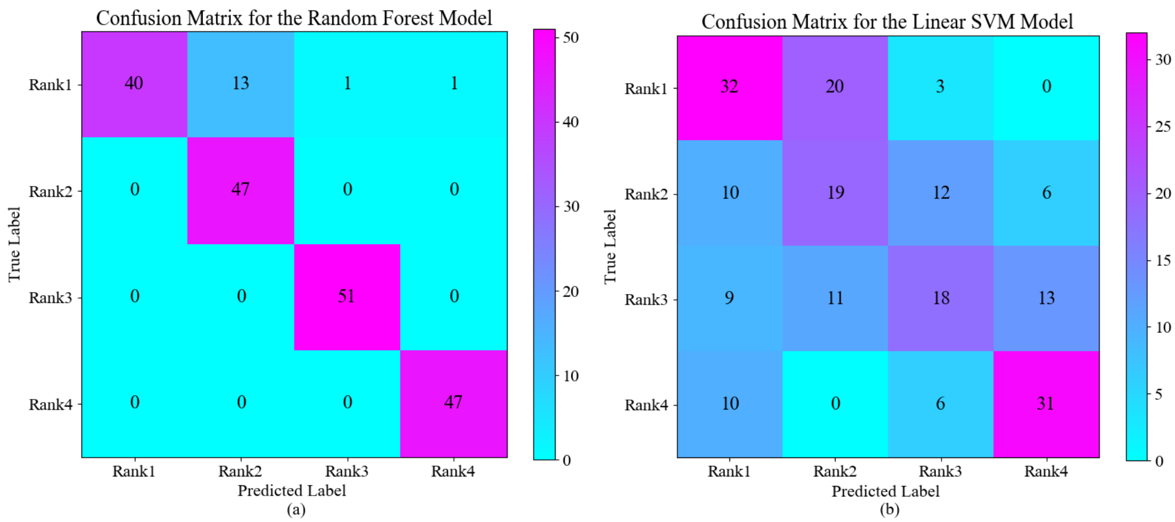


Figure 5.13 Confusion matrices for RF and linear SVM on the test dataset

Figure 5.13a suggests that the classifications of most trucks were successful, except for a small portion of trucks in Rank 1, misclassified as Rank 2. Overall, RF is a strong performer. On the other hand, Figure 5.13b suggests that the linear SVM is a poor performer, with many Rank 4 trucks classified as Rank 1. In addition, the SVM confuses Rank 2 and Rank 3 trucks throughout. These results suggest that RF is the better choice for predicting the performance of snowplow trucks based on the classification accuracy score and confusion matrix.

5.2.2 Feature Importance Analysis

As mentioned earlier, an important feature of RF and linear SVM is their ability to provide information that helps interpret the importance of variables. Understanding the importance of features can help agencies better understand which features are dominant in affecting snowplow truck performance, and assist in the development of prioritization strategies to address fleet deterioration issues. Moreover, it can benefit model construction by filtering out insignificant variables and building a more parsimonious model with improved predictive capabilities. Considering the relatively poor predictive performance of linear SVM detailed above, we limit our results to the use of RF for feature importance analysis. Mean decrease impurity is implemented for each split in RF during the training process, and all features are ranked by the average decrease of impurities across all trees. Table 5.2 presents the feature importance.

Table 5.2 Ranking of variables based on the importance to the performance of snowplow trucks

Variable	Importance Coefficient	Variable	Importance Coefficient
Mi	**0.33**	Func	0.10
Fuel	0.16	Snow	0.09
Year	0.15	Area	0.03
Vol	0.13	Type	0.01

As shown in Table 5.2, working mileage appears to be the most important feature in affecting the performance of snowplow trucks. Fuel consumption and service year are also important factors leading to performance deterioration. AADT, functional classification of roads that trucks serve, as well as annual average snow depth matter to the performance as well. The important takeaway here is that trucks working in different environments can have different performances even with the same work intensity. Also note that loading capacity and area types (rural vs. urban) have an insignificant effect on the major repair times. This means that if we drop these two variables from the dataset, the performance of the model may be only marginally affected.

In broad terms, endogenous features comprise 66% of the total factors and exogenous features account for 34%. In other words, current work intensity of a snowplow truck (e.g., working mileage and service year) is the top priority when considering replacement strategies.

5.2.3 Performance Prediction

One can apply the predictive model to estimate snowplow truck performance over the service time span. Such application can enable effective trend analysis regarding performance deterioration and correspondingly suggest a reasonable level of work intensity. For demonstration purpose, we choose AADT ranging from 0 to 60,000 veh/day to represent different working environments, and working mileage in the range of 3,000 to 15,000 miles/year to represent different work intensities. For simplicity, we assume the fuel consumption rate for Class 8 snowplow trucks is 6 mpg during the entire operation (Hajibabai et al. 2014). Road functional classification and annual snow depth are set as 3.20 and 159.20 ml (average values across all trucks), separately. Lastly, the service region is randomly chosen between urban and rural, and type of load is randomly chosen from types 104, 113, and 168, since these two variables only marginally influence the performance. We predict truck performance for the third, sixth, 10th and 15th years of service separately. Figure 5.14 presents the results of this analysis.

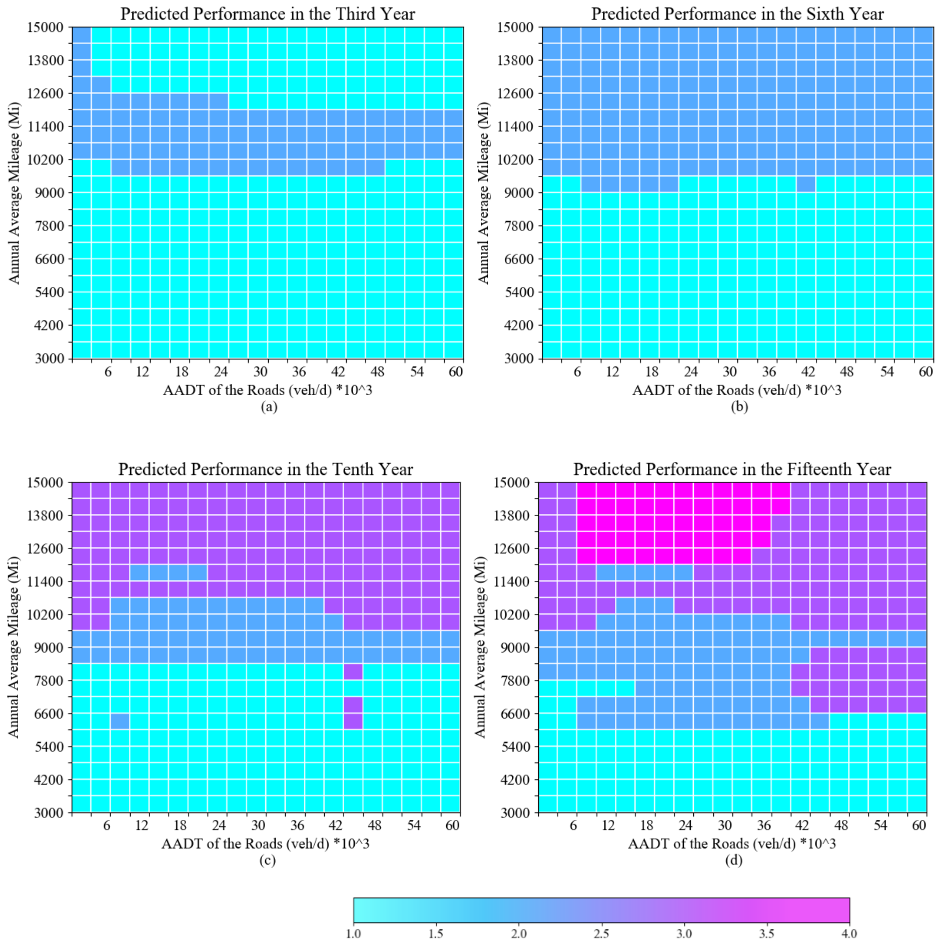


Figure 5.14 Change of major repair ranks of snowplow trucks across different workloads and working environments, where color is distinguished by ranks

As we can see in Figure 5.14, the change of ranks is more sensitive to working mileage than AADT of roads that the truck serves at the same service year, which implies that working mileage is more dominant in affecting performance. In Figure 5.14a, we also notice that trucks working with high annual mileage (above 12,000 miles) are still in Rank 1, while trucks working with annual mileage between 10,000 to 12,000 miles are in Rank 2. We attribute this inaccuracy to the scarcity of training samples with high annual working mileage. Meanwhile, it appears that there is a trade-off between work intensity and performance, where more working mileage leads to more frequent major repairs. In other words, the average working mileage could be appropriately controlled or allocated across trucks to prevent rapid performance deterioration. Specifically, trucks operating less than 6,000 miles annually mainly maintain decent performance even after serving 15 years. Meanwhile, if we set the threshold as 8,000 miles, trucks can service continuously for 10 years while maintaining good performance. If one limits the annual working mileage to 10,000 miles or less, the truck can stay in good condition after serving more than six years. This result can complement life-cycle analysis by adjusting average annual working mileage in terms of the replacement year. In this way, it may lead to less repair costs and more utilization of trucks.

Variation among AADT is not as significant as among working mileage. However, if snowplow truck serves roads with high AADT (over 40,000 veh/day), it should work less frequently than trucks serving other roads to avoid more repairs. More accurate predictions require that one accounts for the variation in the functional classification of roads and annual snow depth as well, since they are also important factors for truck performance.

6. CONCLUSIONS

Snowplow truck performance can affect the road condition and traffic safety during the winter season, especially in regions where the storms are frequent and unpredictable. Thus, the optimizing model is needed to ensure the high efficiency and capability of the snowplow trucks. This project focuses on estimating the optimal life cycle and predicting the operational performance of Class 8 snowplow trucks managed by UDOT.

To achieve the first goal, the cost-benefit model is used to provide a thorough analysis in search of the optimal year for the replacement of the specific types of snowplow trucks. The available maintenance and operation data from 2000 to 2017 justify the effectiveness of the data-driven approach. The results suggest a shorter replacement cycle than what is currently implemented, and provide additional guidance on the procurement, maintenance, and prioritized selection of Class 8 snowplow trucks.

To predict truck performance subsequently, this study employs ML techniques to develop a predictive model by utilizing both endogenous and exogenous features regarding operational performance. We simultaneously train and evaluate RF and linear SVM models for a fleet of 388 individual snowplow trucks. Statistical analysis shows that RF attains an average accuracy of 92% on the test dataset and performs better than linear SVM for the prediction of their repair status. Moreover, we implement mean decrease impurity measures to explore which variables are significant in explaining the deterioration of performance in snowplow activities. The ranking of those features can provide a better understanding of what causes the lowering of performance and can help transportation agencies refine their truck replacement strategy effectively at the micro-level. In addition to feature importance analysis, we apply RF to visualize the change of truck performance with the increase of service years by varying work intensities and working environments. The results suggest a reasonable range of average annual working mileage based on replacement year for preventing quick deterioration of their performance.

Due to data access limitations, we cannot acquire the detailed records pertaining to major repairs. In fact, not all major repairs are directly linked to snowplowing activities, as some repairs may be attributed to the truck's interior structure. One extension for future analysis is to find a method to accurately delineate the repair activities caused by snowplowing so that ML models can be more explanatory to truck performance on snowplow activities.

According to the study developed in this research, the suggested optimal life cycle for Class 8 snowplow trucks managed by UDOT is five years on the statewide level. For trucks serving in some regions (i.e., Salt Lake City and Orem regions), they tend to have longer service spans. However, most Class 8 snowplow trucks were disposed over 13 years of utilization. This longer service span can lead to more frequent major repairs and higher maintenance costs. As a result, UDOT should shorten the average life cycle for Class 8 snowplow trucks to cut down overall expenses.

In fact, a small portion of snowplow trucks can still function with satisfactory operational efficiency over the calculated optimal life cycle. The proposed RF can help UDOT accurately identify the performance of snowplow trucks with a variety of conditions, which can complement the replacement strategy effectively. Moreover, the results indicate that both exogenous and endogenous features regarding snowplow operations can significantly impact truck performance. Hence, we recommend that trucks serving roads with high traffic volumes should function less than trucks serving regular roads to reduce the rate of performance deterioration.

7. REFERENCES

Adams, T. M., M. Danijarsa, T. Martinelli, G. Stanuch, and A. Vonderohe. "Performance Measures for Winter Operations." *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1824: 87-97.

Alpaydin, E. "Introduction and Supervised Learning." In *Introduction to Machine Learning*. MIT press, 2009, pp. 1-44.

Breiman, L. *Classification and Regression Trees*. Routledge, N.Y., 1984.

Breiman, L. "Random Forests." *Machine Learning*, 2001. 45: 5-32.

Bukhsh, Z. A., A. Saeed, I. Stipanovic, and A. G. Doree. Predictive Maintenance Using Tree-based Classification Techniques: A Case of Railway Switches. *Transportation Research Part C: Emerging Technologies*, 2019. 101: 35-54.

Cortes, C., and V. Vapnik. "Support-vector Networks." *Machine Learning*, 1995. 20: 273-297.

Cuelho, E., and D. Kack. *Needs Assessment and Cost/Benefit Analysis of the Roadview Advanced Snowplow Technology System*. California Advanced Highway Maintenance and Construction Technology Research Center (AHMCT) Research Report, 2002.

Dietterich, T. G. "Ensemble Methods in Machine Learning." *International Workshop on Multiple Classifier Systems*. Berlin, Heidelberg, 2000. 1:15.

FHWA, *Road Weather Management Program-Snow and Ice*, Federal Highway Administration, 2016.

Ghobadian, B., H. Rahimi, A. M. Nikbakht, G. Najafi, and T. F. Yusaf. "Diesel Engine Performance and Exhaust Emission Analysis Using Waste Cooking Biodiesel Fuel with An Artificial Neural Network." *Renewable Energy*, 2009. 34: 976-982.

Hajibabai, L., S. M. Nourbakhsh, Y. Ouyang, and F. Peng. "Network Routing of Snowplow Trucks with Resource Replenishment and Plowing Priorities: Formulation, Algorithm, and Application." *Transportation Research Record: Journal of the Transportation Research Board*, 2014. 2440: 16-25.

Hall, S. A., J. S. Kaufman, and T. C. Ricketts. "Defining Urban and Rural Areas in U.S. Epidemiologic Studies." *Journal of Urban Health*, 2006. 83: 162-175.

Kohavi, R. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection." *IJCAI*, 1995. 14: 1137-1145.

Kotsiantis, S. B., D. Kanellopoulos, and P. E. Pintelas. "Data Preprocessing for Supervised Learning." *International Journal of Computer Science*, 2006. 1: 111-117.

Kwon, T. J., and L. Gu. "Modelling of Winter Road Surface Temperature (RST) - A GIS-based Approach." *2017 4th International Conference on Transportation Information and Safety (ICTIS)*, 2017: 10.1109/ICTIS.2017.804782.

Langley, P., and H. A. Simon. "Applications of Machine Learning and Rule Induction." *Communications of the ACM*, 1995. 38: 54-64.

Litman, T. “Transportation Cost and Benefit Analysis: Techniques, Estimates and Implications.” *Transportation Research Record: Journal of the Transportation Research Board*, 1998. 1649: 86-94.

Ma, X., Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang. “Learning Traffic and Images: A Deep Convolutional Neural Network for Large-scale Transportation Network Speed Prediction.” *Sensors*, 2017. 17: 818-834.

Marković, N., S. Milinković, K. S. Tikhonov, and P. Schonfeld. “Analyzing Passenger Train Arrival Delays with Support Vector Regression.” *Transportation Research Part C: Emerging Technologies*, 2015. 56: 251-262.

MesoWest Data, the University of Utah. https://mesowest.utah.edu/. Accessed Feb. 1, 2019.

Perrier, N., A. Langevin, and J. F. Campbell. “A Survey of Models and Algorithms for Winter Road Maintenance. Part I: System Design for Spreading and Plowing.” *Computer & Operations Research*, 2006. 33: 209-238.

Quinlan, J. R. “Induction of Decision Trees.” *Machine Learning*, 1986. 1: 81-106.

Quinlan, J. R. *C4.5: Programs for Machine Learning*. Elsevier, 2014.

Rodriguez, J. D., A. Perez, and J. A. Lozano. “Sensitivity Analysis of K-fold Cross Validation in Prediction Error Estimation.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009. 32: 569-575.

Scheibe, K., S. Nilakanta, and C. Ragsdale. *Decision Support System for Optimized Equipment Turnover*. Transportation Research Board, 2017. InTrans Project: 13-348.

Scikit-learn: “Machine learning in Python.” https://scikit-learn.org/stable/. Accessed Feb. 1, 2019.

Sun, Y., Z. Jiang, J. Gu, M. Zhou, Y. Li, and L. Zhang. “Analyzing High Speed Rail Passengers’ Train Choices Based on New Online Booking Data in China.” *Transportation Research Part C: Emerging Technologies*, 2018. 97: 96-113.

Swanson, L. “Linking Maintenance Strategies to Performance.” *International Journal of Production Economics*, 2001. 70: 237:244.

The Census Transportation Planning Products Program. http://data5.ctpp.transportation.org

/ctpp/Browse/browsetables.aspx. Accessed Feb. 1, 2019.

Van Hulse, J., T. M. Khoshgoftaar, and A. Napolitano. “Experimental Perspectives on Learning from Imbalanced Data.” *Proceedings of the 24th International Conference on Machine Learning. ACM*, 2007. 935-943.

Wyrick, D. A., and S. Erquicia. *Fleet Asset Life Cycle Costing with Intelligent Vehicles*. University of Minnesota Center for Transportation Studies, Retrieved from the University of Minnesota Digital Conservancy, 2008.