Aalto University
MASTER'S THESIS 2021

Decision Support Model for Fleet Maintenance based on Vehicle Quality Incentives for Bus Operators in Public Transportation

Oskari Kivinen

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo, 6 June 2021

Supervisor:	Professor Antti Punkka
Advisors:	M. Sc. Petri Auno, Nobina Oy
	B.Eng. Marko Lonnakko, Nobina Oy

Aalto University
School of Science
Master's Programme in Mathematics and Operations Research

Aalto University
Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Oskari Kivinen

Title
Decision Support Model for Fleet Maintenance based on Vehicle Quality Incentives for Bus Operators in Public Transportation

School School of Science

Master's program Mathematics and Operations Research

Major Systems and Operations Research **Code** SCI3055

Supervisor Professor Antti Punkka

Advisors M.Sc. Petri Auno & B.Eng. Marko Lonnakko, Nobina Oy

Level Master's thesis **Date** 6 June 2021 **Pages** iv+36 **Language** English

Public transportation is often organized and operated by separate organizations. The public authority in Helsinki, HSL, is responsible for organizing the time schedules, route planning etc. The operation of the bus transportation is outsourced to bus operators, such as Nobina. HSL has set different incentive and sanction systems to control operation to a favourable way. One of the incentive systems is related to the visible quality of the fleet in operation. The quality is monitored on contract level during inspection seasons by giving inspection points on specific faults, and the performance is measured by counting inspection points received per inspection, with lower amounts resulting to higher bonuses.

This thesis develops a decision support tool to target limited maintenance actions to accumulate more quality incentive bonuses. The objective was to develop a process to find a better policy (or maintenance order), and create a tool to determine the best order at any point in time. The system, i.e. bus operation in the context of incentive bonuses, was modeled as a Markov Decision Process (MDP). The current faults, inspection points, and inspections formed the states, maintenance procedures the actions, and the incentive bonus as the payoff from the MDP. The policies were tested by simulation to determine the expected payoff for each policy.

The results suggest that it is possible to create better policies in the process. A sensitivity analysis revealed that preventing faults being caught in an inspection has the highest impact on the received incentive bonus. Further development of the model could change the results, as some key parts were forced to be left out, such as the cost of actions. In addition, policy optimization could likely further improve the gain of incentive bonuses.

Keywords Markov Decision Process, public transport

Aalto-yliopisto

Tiivistelmä

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tekijä
Oskari Kivinen

Työn nimi
Päätöksentekotyökalu kaluston huoltoon perustuen ajoneuvon kuntoon liittyviin kannustimiin julkisessa liikenteessä

Korkeakoulu Perustieteiden korkeakoulu

Maisteriohjelma Mathematics and Operations Research

Pääaine Systems and Operations Research **Koodi** SCI3055

Valvoja Professori Antti Punkka

Ohjaajat FM Petri Auno & Tekn. ins. Marko Lonnakko, Nobina Oy

Työn laji Diplomityö **Päiväys** 6.6.2021 **Sivuja** iv+36 **Kieli** englanti

Tiivistelmä

Julkinen joukkoliikenne on usein järjestetty ja liikennöity eri organisaatioiden toimesta. Julkinen taho Helsingissä – HSL (Helsingin seudun liikenne) – on vastuussa mm. liikenteen aikatauluista ja reitityksestä. Bussien liikennöinti on ulkoistettu liikennöitsijöille, kuten Nobinalle. HSL on asettanut erilaisia kannustin- ja sanktiojärjestelmiä hallitakseen liikennöinnin toteuttamista toivottuun suuntaan. Yksi kannustinjärjestelmistä liittyy käytetyn bussikaluston nähtävissä olevaan laatuun. Laatua valvotaan sopimustasolla tarkastuskausien aikana antamalla tarkastuspisteitä tietyistä vioista, ja suoriutumista mitataan laskemalla keskimääräiset tarkastuspisteet tarkastusta kohden. Alhaisempi lukema johtaa suurempaan kannustimeen.

Tässä työssä kehitetään päätöksentukityökalu rajallisten huoltotoimenpiteiden kohdentamiseen isompien laatukannustimien kartuttamiseksi. Tavoitteena oli löytää parempi huoltokäytäntö (tai huoltojärjestys) ja luoda työkalu parhaimman järjestyksen määrittämiseen minä ajanhetkenä tahansa. Systeemi, eli bussien liikennöinti kannustimien kontekstissa, mallinnettiin Markovin päätöksentekoprosessina. Olemassaolevat viat, tarkastuspisteet ja tarkastukset muodostivat prosessin tilat, huoltotoimenpiteet prosessin toiminnat ja kannustimet prosessin palkinnon. Huoltokäytännöt testattiin simuloinnin kautta, jotta odotettu tuotto voitiin määrittää kullekin käytännölle.

Tulokset kielivät mahdollisuudesta luoda parempia käytäntöjä prosessiin. Herkkyysanalyysi paljasti, että vikojen kiinnijäämisen estämisellä tarkastuksissa on suurin vaikutus saatuihin kannustimiin. Mallin lisäkehitys voi muuttaa tuloksia, sillä joitain pääkohtia jouduttiin jättämään mallista pois, kuten toimintojen kustannukset. Tämän lisäksi käytäntöjen optimointi voisi todennäköisesti edelleen kasvattaa saatujen kannustimien määrää.

Avainsanat Markovin päätöksentekoprosessi, joukkoliikenne

Contents

1. Introduction

Public transport has been used in city areas from early 19th century (Parks 2017), and with the megatrend of urbanization, the utilization of public transport has been estimated to increase heavily (PwC 2020). The introduction of public transport has been a key factor in the expansion of cities to sub-urban areas, hence offering a solution to mobility in areas with ever-increasing population density. For the climate change, public transport introduces a way to decrease emissions, as urban mobility is a major source of transport emissions (European Comission 2020).

Government or municipal subsidies are a common way to finance public transport due to the collective benefits it provides. Some cities have a public authority to design and manage the ticket systems, timetables and route planning, such as SL in Stockholm and HSL in Helsinki (Jansson and Pyddoke 2010; The Board of HSL 2019). In Helsinki, the authority is owned by the cities and municipalities that participate in the transport with the expenses divided by the amount of traffic each city has. The transport itself is then managed by separate operators, such as VR for the commuter trains (Pääkaupunkiseudun Kalusto Oy 2020) and Nobina Oy for a share of the buses (Kuukankorpi 2020). Nobina is the Nordic region's largest and most experienced public transport service provider. Nobina is the largest public transport service provider in the Helsinki region and thus also one of the country's major players.

Since the authority—HSL in Helsinki's case—is responsible for the customer experience for the passengers, the contracts with the bus operators have both incentives and sanctions to guide the operators to perform well on selected criteria, such as punctuality, cancellations and the quality of the used fleet. The quality of the fleet can be divided as the features of the fleet, such as the model and age, and the visible quality, such as cleanliness and integrity. HSL has set incentives to improve the visible quality by giving bonuses of up to 1% of accumulated revenue in a bonus season for operators who maintain their fleet in top condition. HSL monitors the quality with subcontracted inspectors, who perform random inspections with ongoing transportation and give inspection points on the emerged faults for each bonus season. Nobina has an incentive program that aims to increase the profit gained by the offered incentives.

This thesis develops a decision support tool for the quality incentive program at Nobina Oy. Since the resources of maintenance are limited, the tool helps on focusing the resources more efficiently by accumulating more incentive bonuses with a good maintenance policy. In general, a policy determines guidelines, from which actions are based on. In Nobina's case, a policy determines where the fixes should be focused on, i.e., the order of contracts to focus on. The objective of the thesis was to develop and test different policies to discover a more lucrative way of repairing faults compared to the current one.

The system that the tool is applied can be modeled as a Markov Decision Process (MDP) to consider the present situation in a bonus season and model when and what type of new faults appear, are inspected, and are repaired. The MDP uses states to represent the inspection points per contract and current faults in contracts based on the present knowledge of which buses have faults and in which contracts they are used. In addition, the MDP considers actions to perform on the MDP states, namely the fixes done to repair faults, and a reward function to measure the effect of the actions. The reward function is determined by the paid bonus at the end of a bonus season, which is based on the amount of inspection points and inspections on contracts, of which only the former can be controlled by Nobina's actions. The MDP advances with a time step of one day and actions are executed based on the effective policy.

The testing was performed by simulating the policies within the system. Each policy produces its own MDP due to different sets of actions leading to different states and eventually different rewards at the end of the process. The system contains four different random variables, i.e. new inspections, new faults, new points from the more severe faults, and the amount of repair actions, which were modeled as Poisson point processes. The realizations of the random variables were the same between policies for each simulation.

As the maximum bonus offered is one percent in HSL region, and typical revenues of bus operators range from tens of millions to a hundred million, the theoretical total benefit for a company is in the order of magnitude of a million euros. Currently, Nobina receives approximately a third of the theoretical maximum. The considerable amount of bonus left motivates for further research on improving current protocols of maintenance actions.

This thesis is structured as follows. Chapter 2 reviews relevant literature of service quality, discrete finite MDP, and Poisson point process. Chapter 3 presents the development of the model by first describing the system, then continuing with the time evolution of the modeled system, and finally describing the estimation of parameters used in the model. Chapter 4 presents the development of different policies and how the model is simulated, followed by the performance of the policies and a sensitivity analysis with regards to the estimated parameters. Then, the model and results are reviewed in Chapter 5. Finally, conclusions on the thesis are presented in Chapter 6.

2. Background

This Chapter introduces the reader to the problem, beginning with the operating environment and the incentive bonuses in Section 2.1. Next, in Section 2.2 the subject of Markov Decision Process (MDP) is expatiated, which was used to build the decision support tool. Finally, some theory of Poisson point processes is reviewed in Section 2.3.

2.1 Service quality in public transportation

Service quality can be determined by three dimensions; the functional quality of the process, the technical quality of the encounter, and the corporate image (Thai et al. 2014). While there exists other definitions (e.g. Parasuraman et al. 1988), the three dimensions offer a comprehensive approach, and thus, are sufficient to describe perceived quality in public transportation.

The functional quality includes the service process, which in the case of public transportation is displayed as e.g. the frequency and punctuality of departures. In addition, the coverage of the transportation network, connections of routes, and methods of payments are considered as functionalities in the service process. In the context of this thesis, HSL — the public transport authority in Helsinki region — is responsible for most of the service process, i.e. the functional quality of the public transportation, since it plans the coverage, connections, and timetables of the transportation, handles the marketing of the service, and hosts the payment system of the service. However, since HSL does not operate the transport itself, it monitors the punctuality of different operators and has set both incentives and sanctions to ensure transport operators fulfill functional quality standards of the transportation which are determined in contracts. The contracts are agreements between the authority and an operator about some proportion of the total transportation with some revenue associated to it. In general, several bus lines are included in one contract, where a bus line is a predetermined route for a bus to drive within the timetables.

On the other hand, HSL as an authority plays a unique role to the customer of the transportation; the whole service is highly associated with the authority itself, although the execution and delivery of the service is mostly dependent on the operators. Whenever a customer uses the websites of the service, pays a ticket, gives feedback on the service, or even climbs on board of a vehicle, the colors and logo of HSL are everywhere. The result is that the corporate image lies essentially fully on HSL, in which it does relevantly well; 76% of passengers were satisfied for the public transportation overall (HSL 2020). However, all of the feedback is forwarded to the target of the feedback, and customer satisfaction is one source of incentives and sanctions. Thus, operators pursue for a good corporate image by e.g. being reliable and responsible. In addition, HSL, media, public opinion, and prospective employees are needed for maintaining and creating business.

Finally, the technical quality includes the visible quality of the service, e.g. how well the service is functioning or how the delivery is perceived. While HSL manages the IT side of the service, the actual transportation is operators' responsibility, which is the largest factor in the service of what the customer perceives whilst using the service. The technical quality ranges from the customer service of the driver to the smoothness of the transport ride and from perceived cleanliness to the intactness of the bus.

As stated earlier, HSL desires to maintain the overall quality of the service since it holds the largest stake in terms of service quality as perceived by the customers. As a consequence, both sanctions and incentives are set to the technical side of the service, since HSL has no other control over how operators deliver the service. One of the incentives is related to the cleanliness and the visible shape of the vehicles. The way these incentives work is that HSL monitors the cleanliness and shape with subcontracted inspectors riding the vehicles during operation. The inspectors report any defects or faults they encounter to a monitoring system. Most defects are rated with an inspection point system, e.g. 10 inspection points for a minor fault and 90 inspection points for an egregious fault. The operators then act based on these reports to fix the faults accordingly.

The inspections used to monitor the quality are executed during inspection periods of bonus seasons. Bonus seasons are always the first and second half of the year, while inspection periods last approximately four months in both spring and fall. The inspection points are used to calculate the bonus points, which in turn determine the amount of paid bonus in a bonus season for each contract. Bonus points for a contract are the average inspection points per inspection during a bonus season. The incentive bonus is determined with a step-wise function, with the bonus being either 1%, 0.5%, or 0% of the accumulated revenue of the contract during the half of the year in which the ongoing bonus season is.

There are two different threshold systems, the old and the new system, to determine the amount of bonus each contract has. The old system is still active, but any new contracts are formed with the new system. For the old system, the contracts are ranked based on the amount of bonus points they have, where the lowest score receives the best rank. The contracts are divided roughly into thirds based on their revenue. The best third with the lowest amount of points receive the maximum bonus of 1%, the second third a 0.5% bonus and the last third receives no bonus. Consequently, the threshold levels in the old systems are dynamic and vary from season to season. For the new system, the bonus levels are pre-determined. If the contract has 10 bonus points or fewer, it receives the best bonus. If it has more than 10 points but under 25 bonus points or exactly 25 points, the contract receives 0.5% bonus. Otherwise, there will be no incentive bonus for the contract. Based on historical threshold levels, the new system rewards more equally on excellent service, but is more strict on the second threshold levels. The common factor between the systems is that a higher amount of bonus points leads to a lower incentive bonus.

2.2 Discrete finite Markov Decision Process

Markov Decision Process (MDP), introduced in the 1950's (Bellman 1957), is a mathematical framework used in decision making. Sutton and Barto (2018) described it as a discrete stochastic version of the optimal control problem in dynamic programming. The MDP is an extension of *Markov chains* with the addition of decisions that affect the state transitions. A Markov chain describes a sequence or a chain of random variables, or states, that all possess the Markov property. The *Markov property* refers to the memoryless property of a stochastic process (Markov 1954). In mathematical terms, the Markov property can be written as

$$P(X = s_{n+1}|X_n = s_n, ..., X_0 = s_0) = P(X_{n+1} = s_{n+1}|X_n = s_n),$$

where X_n denotes a random variable at time t_n, and $s_n \in S$ denotes a state in state space S, which consists of all possible states (Durrett 2019). The state space is a measurable space $(S, \mathbf{S})$, where $\mathbf{S}$ denotes the algebras or operations related to S. A function $q : S \times \mathbf{S} \to \mathbb{R}$ is a transition probability of a Markov chain if it satisfies the following conditions:

1. For each $x \in S, O \to q(x, O)$ is a probability measure on $(S, \mathbf{S})$.

2. For each $O \in \mathbf{S}, x \to q(x, O)$ is a measurable function.

Moreover, we say X_n is a Markov chain with a transition probability q if

$$P(X_{n+1} \in B|\mathcal{F}_n) = q(X_n|B),$$

where $\mathcal{F}_n$ is a nonempty collection of subsets of all possible outcomes related to the Markov chain, and B denotes the set of possible states of Markov chain (Durrett 2019).

The extensions the MDP introduces to a Markov chain are rewards and actions. The agent in control of the actions interacts with the environment at discrete time steps. Based on the state $S_t \in S$ at time t the agent selects an action $A_t \in A(s), s \in S$. Based on the action, the state moves to S_{t+1} and a reward R_{t+1} is received. The goal for the agent is to maximize the cumulative reward (or if the reward R is negative then minimize the cumulative costs) denoted by G_t. In many cases, rewards are discounted to take the time value into account. Total reward at time t is calculated as

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + ... = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1},$$

where γ is the discount factor for the time corresponding the time value. The first terms can be rearranged to find a recursive manner on the total reward:

$$
\begin{aligned}
G_t &= R_{t+1} + \gamma \left(R_{t+2} + \gamma R_{t+3} + ...\right) \\
&= R_{t+1} + \gamma G_{t+1}.
\end{aligned} \tag{2.1}
$$

Both S_t and R_t are random variables that follow some probability distribution. The variables depend only on the preceding values of S_{t-1} and A_{t-1}. The dynamics of the MDP can be defined as

$$p(s', r|s, a) \doteq \mathbf{Pr}\{S_t = s', R_t = r|S_{t-1} = s, A_{t-1} = a\},$$

for all $s', s \in S, r \in R,$ and $a \in A(s)$ (Sutton and Barto 2018).

The function p fully describes the probability distributions related to states and rewards, and, as the MDP should, it holds the Markov property. In addition, p can be used to calculate *state-transition probabilities*:

$$p(s'|s, a) \doteq \mathbf{Pr}\{S_t = s'|S_{t-1} = s, A_{t-1} = a\} = \sum_{r \in R} p(s', r|s, a).$$

To support the decisions regarding the choice of actions, a *policy* π is determined by the agent. The policy should be a robust way to perform actions on each state. Moreover, the policy should only depend on the current state and not the history due to the Markov property:

$$\pi(a|s) = \mathbf{Pr}\{A_t = a|S_t = s\}.$$

Following a policy π on a state s returns a deterministic reward, if only the actions impact the state transition, i.e. no impact from environment occurs. The objective is to maximize the reward (or minimize the costs) by selecting the best policy π among a group of policies. *State-value function* denoted by v_π is used to calculate the expected return of the process, and is written as

$$v_\pi(s) \doteq \mathbb{E}_\pi[G_t|S_t = s] = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s \right].$$

Similarly to the total reward (2.1), the state-value function too can be written recursively:

$$\begin{aligned}
v_\pi(s) &\doteq \mathbb{E}_\pi[G_t|S_t = s] \\
&= \mathbb{E}_\pi[R_t + \gamma G_{t+1}|S_t = s] \\
&= \mathbb{E}_\pi[R_t + \gamma v_\pi(S_{t+1})|S_t = s] \\
&= \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a) \left[r + \gamma v_\pi(s')\right] \forall\ s \in S.
\end{aligned}$$

The last row fulfils the *Bellman equation* for v_π, which is considered in dynamic programming as a necessary condition for optimality (Dixit and Sherrerd 1990).

While state-value function $v_\pi(s)$ shows the value of a state s by following a policy π, an *action-value function* $q_\pi(s, a)$ shows the value of taking an action a on a state s under a policy π, as

$$\begin{aligned}
q_\pi(s, a) &\doteq \mathbb{E}_\pi[R_t + \gamma v_\pi(S_{t+1})|S_t = s, A_t = a] \\
&= \sum_{s',r} p(s', r|s, a) \left[r + \gamma q_\pi(s', a')\right] \forall\ s \in S.
\end{aligned}$$

The value of a policy may be calculated using the state-value function, since the sum of the values of all states is the result of the chosen actions determined by policy π. This allows a straightforward comparison between different policies, where the largest value describes the best policy. A policy π can be further improved with the action-value function, as it describes if any action a' yields a larger value than the action a given by the policy π.

Due to flexibility of the framework for problems in the world, the MDP is applied in numerous different situations. The applications include e.g. herd management in animal production (Kristensen 1996), optimal electricity supply bidding (Song et al. 2000), dynamic pricing models (Aviv and Pazgal 2005), maintenance scheduling (Leppinen 2020), and finance and investment models (White 1993). In addition, the field of machine learning utilizes the MDP, since reinforcement learning is heavily inspired by it (Sutton and Barto 2018).

2.3 Poisson point process

Several phenomena in the nature can be modeled as stochastic processes, with realizations appearing as point events in space or time. Point events in space include e.g. the distance between stars, the location of molecules in gas, and the distribution of blood cells in a sample (Guttorp and Thorarinsdottir 2012), whilst point events in time include e.g. frequency of earthquakes, radioactive decay, and cars passing a reference point in a road (Cha 2018). The point process is called a Poisson point process, or Poisson process, if the events follow a Poisson distribution (Cha 2018). Especially events with the characteristics of vast amount of possible realizations, but a relatively low probability of occurring, such as the location of a molecule in gas, are often Poisson distributed. Poisson Processes were the foundation from which Markov processes were developed (Guttorp and Thorarinsdottir 2012; Itô 2020).

Poisson distribution is a discrete probability distribution, expressing the density of independent events occurring. For example, Stirzaker (2000) derives a general probability mass function (PMF) for the Poisson distribution from the binomial distribution. The PMF of Poisson distribution can be written as

$$\Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{2.2}$$

where λ depicts the frequency of the events, and k is the number of events. In fact, λ is both the expected value and the variance of X. The expected value can be derived from the rate of occurrence for the events as $\lambda = rt$, where r is the number of events per unit of time.

The behavior of any phenomenon in the nature that follows a Poisson distribution can be predicted with the probabilities given by the PMF (2.2). The probabilities express the likelihood of observing the events in the future. The probabilities can be determined given that λ can be calculated by observing r. The Poisson process may be simulated using the calculated probabilities. A common and an easy way to simulate the Poisson distribution is to use the cumulative distribution function (CDF), and with a uniformly distributed random variable between zero and one determine the amount of events from the CDF (Date 2019).

3. Model development

This Chapter clarifies the modeled system, which is the operation of the buses in the context of the quality incentive program. In Section 3.1, the key concepts are introduced to understand how the system operates, and explanations are given on why some approximations and simplifications are necessary in the model of the system. Section 3.2 further explains the evolution of the modeled system both in general and with an example of a miniature system. In Section 3.3 the parameters and the basis for their selection are reviewed in detail.

3.1 Description of the system

This thesis models the operation of buses in the context of the quality incentive program using a discrete-time Markov Decision Process (MDP), including states, actions, and a reward at the last state of the MDP. The MDP is used to determine the best policy π among a set of feasible policies to perform certain actions on different states to produce maximal incentive bonuses from a bonus season. The decision period, i.e. the time interval of the process from time step t to $t+1$, is one day, and each event or action occurs or is performed during each period. Both events and actions change the state of the system and ultimately affect the expected incentive bonus at the end of the bonus season. For example, new faults and inspections are considered events, while fixes are considered actions.

In the context of this thesis, faults are considered as anything that can be perceived and which decline the quality of the service for a passenger. They include any dysfunctional items, such as a non-operational stop-button, or items that cause discomfort, such as a dirty seat. Faults occur due to accidents in traffic, vandalism, or neglect for tidiness by the passengers. The nature of the faults is that they are unpredictable and thus arise randomly.

The faults that are inspected are given inspection points based on a separate inspection point table, which is distributed by HSL to all operators. A part of the whole table is shown in Table 3.1. Each fault belongs to a fault category describing the location of a fault, such as doors, stop-buttons, or seat cleanliness, totaling to 17 different categories.

Table 3.1. Faults are given inspection points based on their fault category and the severity of the fault if caught in an inspection. This Table displays a part of the whole point table that HSL has distributed to the operators.

Severity	Doors	Stop-buttons	Seat cleanliness	Front plate
Dangerous	90	-	60	-
Must be fixed	60	20	40	60
Minor	30	10	20	30

Each fault in every category is given a severity status based on how much discomfort each fault produces. In addition to faults giving inspection points, the faults may incur a sanction if not repaired within an agreed deadline. Dangerous faults may cause physical harm to passengers and need to be repaired within a day, or the vehicle is not allowed to operate anymore. Faults that must be fixed have a deadline of six months to be repaired before sanctions are given. These faults are not physically dangerous, but are thought to significantly lower the overall attractiveness of traveling. Finally, minor faults have mainly a mild unaesthetic effect and do not incur any sanctions if not repaired, although the same faults may be caught in inspections, consequently lowering the incentive bonus.

Ideally, all faults would be fixed whenever one is observed either by self-monitoring the vehicles or through inspections. Typically, bus operators have their own maintenance to fix all kinds of faults that may appear in the vehicles. However, maintenance has several limitations, which come down to capacity, cost, and vehicle type restrictions on contracts. More often than not, there is not enough room, staff or time, i.e. capacity, to fix every fault that is observed. Thus, it is important to have a clear priority, a maintenance order, for all fixes. In this thesis, different policies basically control the maintenance order to maximize the incentive bonus by reducing inspection points on future inspections in key contracts.

Maintenance costs can be roughly divided as labor and material costs. While labor costs are relatively constant between fault categories, the material costs vary greatly depending on where the fault is located and how severe the fault is. For example, fixing a stop-button may cost approximately 50 euros, where as fixing a scratch on the body of the bus may cost around a thousand euros. Despite the large discrepancy between the costs of fault categories, the inspection points do not differ as much. Therefore, it would make sense to control the maintenance order by those faults that have the best impact on bonus per cost ratio. Unfortunately, the inspection data gives insufficient information to draw conclusions about the approximate costs, as many faults have cost ranges from 300 to 1000 euros. In addition, it is possible to have many faults in a small area which could be fixed for a price of one repair action. This is, however, also undisclosed in the inspection data, ergo credible approximations are impossible to conduct. Similarly, the vehicle restrictions on contracts is a complex topic, since whenever a bus is under maintenance, it requires a reserve bus on the contracts to be used. The vehicle restrictions are difficult to simulate due to imperfect data, and easily leads to infeasible combinations of reserve buses in traffic, which incurs a sanction. Both costs and vehicle restrictions are omitted from the system considered in the thesis to simplify the model.

The minor faults clearly differ from the two other types of faults. The minor faults receive a lower priority in the maintenance order both in terms of received sanctions – none for minor faults – and lost incentive bonus due to lower amount of inspection points received. These faults often remain unrepaired for extended periods and accumulate inspection points from several inspections. In general, Nobina receives approximately half of their inspection points from minor faults. In the context of the thesis, the action-value function would simulate the cost of fixes, thus yielding a direct negative impact. However, due to an uncertain manner of the costs, it is extremely difficult to estimate a reliable cost for any repair actions. Moreover, all of the existing faults are sought to be fixed which means that the costs would incur at some point anyway, and since there is no discounting, there is an indifference between fixing a fault now versus later. Thus, minor faults are the focus of this thesis and only they are modeled in the system as faults to be repaired, as the other types of faults are trivial since they will be repaired nonetheless as soon as possible.

Ideally, all of the emerged faults are to be noticed before they are caught in an inspection, enabling actions to prevent more points being gathered from inspections either by repairing the fault or switching the bus to another one from reserve or another contract. However, monitoring the faults is imperfect due to lack of resources and in some cases secretion from those who have caused the fault. Consequently, some faults will only be noticed after they have been caught in an inspection.

Switching buses might prove useful if there is a contract which has room for inspection points due to the stepwise nature of the paid incentive bonus function. However, as is the case of using spare buses during maintenance, switching buses between contracts has many restrictions, as the vehicle type used in contracts is regulated and controlled via sanctions. In addition, monitoring available buses requires manual work to track where each bus is located at any given time. Thus, feasible switches are difficult to model due to complex and strict rules outside the scope of this thesis, which is why bus switches between contracts are omitted from the modeled system.

Inspections are performed by a subcontracted service. HSL has set target amounts of inspections for each bus operator based on the registered fleet size each operator owns. Inspectors choose relatively randomly different bus lines, and board seemingly random buses. They investigate the bus both inside and outside. Since it is their single task to observe any and all faults, it is fair to assume that they will find any faults that exist in the bus during the inspection. The incentive bonus is paid based on the performance of operators in inspections. Each contract has its own bonus points, which is calculated as the average inspection points per inspection. Thus, bonus points can be both improved or worsened in the span of an inspection season. For a more detailed explanation of the bonus point system, refer to Section 2.1.

3.2 Time evolution of the modeled system

The system is modeled on a day-to-day basis for the duration of an inspection season, giving the limits for a discrete time step $t \in [0, T]$, where T is the last day of the inspection season. The system consists of I fault categories and J contracts. I is determined by HSL, and it is possible to change it between bonus seasons. J depends on the operator and how well they fare against the competition in tenders. A state $x(t)$ of the modeled system described above is determined by the total inspection points p_j and inspections c_j per contract $j \in [1, J]$, and the existing minor faults v_{ij} per fault category $i \in [1, I]$ per contract j. The inspection points and inspections per contract at time T determine the bonus threshold in which the contract is at the end of the bonus season. The faults per fault category – or fault type – per contract are based on bus level data, which is consolidated to contract level, so the system excludes information of faults per buses.

The first state at $t = 0$ is based on realized data, such as the effective driving schedule and the existing faults per buses. New faults and repair actions follow a Poisson point process, and they are added on top of the known faults. The faults affect the expected inspection points gained through inspections, which also follow a Poisson point process. The state can be written as

$$x(t) = \begin{bmatrix} v_{11} & \dots & v_{1J} \\ \vdots & \ddots & \vdots \\ v_{I1} & \dots & v_{IJ} \\ c_1 & \dots & c_J \\ p_1 & \dots & p_J \end{bmatrix}.$$

For the purpose of this Section, let us construct an example system. Suppose $I = 4$ and $J = 3$ with a bus capacity b_c of three buses for each contract, which means that there are three buses used each day for each contract. b_c is also used in the inspection point calculations later in this Section. The fault categories are faults on doors, stop-buttons, seat cleanliness, and the front plate of the bus, which displays the line number and destination of the bus. The example state x_e is then

$$x_e(t) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 3 & 0 \\ 4 & 3 & 2 \\ 50 & 130 & 10 \end{bmatrix}.$$

The system evolves in time through transitions of states. The state transitions are affected by the chosen actions, since repairing a fault changes the faults in the state. Due to limited resources, the workshops are unable to repair all the existing faults, and due to the various nature of the faults, the amount of repair actions in a day is not constant. Instead, the action count can be modeled as a Poisson point process. The actions are denoted as

$$u(t) = \begin{bmatrix} u_{11} & \dots & u_{1J} \\ \vdots & \ddots & \vdots \\ u_{I1} & \dots & u_{IJ} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix},$$

where u_{ij} is the change of the amount of faults of type i in contract j due to fixes done at time t. The example state repair actions could be focused on contract $j = 2$, giving an example realization of actions as

$$u_e(t) = \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

After the selection of actions, let us consider the state transitions from time t to $t+1$. For the duration of the time step, all the actions are implemented, in addition to the additional faults, inspections, and inspection points being added to the state, all of which follow their own Poisson point processes. The new faults are denoted as

$$w_f(t) = \begin{bmatrix} V_{11} & \dots & V_{1J} \\ \vdots & \ddots & \vdots \\ V_{I1} & \dots & V_{IJ} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix},$$

where V_{ij} is the amount of faults of type i for contract j. V_{ij} also follows the Poisson process. The example state could face new faults e.g. as

$$w_{f,e}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

New inspections are modeled as

$$w_c(t) = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ C_1 & \dots & C_J \\ 0 & \dots & 0 \end{bmatrix},$$

where C_j is the amount of new inspections on contract j. The example state could receive inspections as

$$w_{c,e}(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The inspection points are dependent on inspections occurring in a state transition. For example, contract $j = 1$ did not receive any inspections in the time step, thus gaining zero inspection points despite having some faults. Since the assumption is that all faults are found on each inspection, the amount of inspection points gained from the minor faults is known for a bus, denoted as m_j. However, since the system in this thesis omits bus level data, the amount of points from minor faults is assumed as if the minor faults are spread evenly on the buses used for the day. The inspection on contract $j = 2$ in the example state would then observe $1/3$ minor faults of type $i = 3$, the seat cleanliness, accumulating $20/3$ inspection points as per Table 3.1. In addition, faults with a severity status of dangerous or must be fixed, i.e. the more severe faults, are modeled as if there are none at any time t, but some may appear during the state transition and get caught in an inspection, accumulating some inspection points. After the more severe faults have been observed through inspections, they are repaired instantly, and do not cause more points for later state transitions. Points gathered in this way from one inspection are denoted as s_j and can be modeled as a Poisson point process. The total inspection points from the more severe faults are simplified such that each new inspection causes s_j points, i.e. the total inspection points from the more severe faults can be calculated as $S_j = C_j s_j$. Thus, the new inspection points are denoted as

$$w_p(t) = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ 0 & \dots & 0 \\ m_1 + S_1 & \dots & m_J + S_J \end{bmatrix},$$

where $m_j = C_j \sum_{i=1}^{I} n_i(v_{ij} + u_{ij} + V_{ij}/2)/b_{c,j}$ are the inspection points from minor faults for contract j. Here, n_i is the point count for a fault of type i and $b_{c,j}$ is the bus capacity for contract j. Notice, that the model assumes that only half of the new faults occur before any inspection and the other half occurs after all inspections. This assumption is based on the dispersion of inspections throughout a day, where the average of the times of inspections is close to noon. To demonstrate the calculation for the example state, the new minor inspection points for contract $j = 2$ are

$$\begin{aligned} m_2 &= 1 \cdot \left(30\frac{1-1+0}{3} + 10\frac{2-1+0}{3} + 20\frac{2-1+1/2}{3} + 30\frac{3-2+0}{3} \right) \\ &= 0 + \frac{10}{3} + 10 + 10 \approx 23.3. \end{aligned}$$

Similarly, the minor fault points can be calculated for contract $j = 3$ as $m_3 = 10\frac{1/2}{3} \approx 1.7$. In addition, an example realization of the inspection points gained from the more severe faults is added, giving the total new inspection points to the example state as

$$w_{p,e}(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 23.3+15 & 1.7+10 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 38.3 & 11.7 \end{bmatrix}.$$

The evolution of the system can be formulated by combining the different parts above to a new function f, described as

$$\begin{aligned} x(t+1) &= f(x(t), u(t), w_f(t), w_c(t), w_p(t)) \\ &= x(t) + u(t) + w_f(t) + w_c(t) + w_p(t) \\ &= \begin{bmatrix} v_{11} & \dots & v_{1J} \\ \vdots & \ddots & \vdots \\ v_{I1} & \dots & v_{IJ} \\ c_1 & \dots & c_J \\ p_1 & \dots & p_J \end{bmatrix} + \begin{bmatrix} u_{11} & \dots & u_{1J} \\ \vdots & \ddots & \vdots \\ u_{I1} & \dots & u_{IJ} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} V_{11} & \dots & V_{1J} \\ \vdots & \ddots & \vdots \\ V_{I1} & \dots & V_{IJ} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix} + \\ & \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ C_1 & \dots & C_J \\ 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ 0 & \dots & 0 \\ m_1 + S_1 & \dots & m_J + S_J \end{bmatrix}. \end{aligned}$$

The state transition for the example state summarizes to

$$x_e(t+1) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 3 & 0 \\ 4 & 3 & 2 \\ 50 & 130 & 10 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 38.3 & 11.7 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \\ 1 & 1 & 0 \\ 4 & 4 & 3 \\ 50 & 168.3 & 21.7 \end{bmatrix}.$$

The objective is to maximize the incentive bonuses gained from contracts at the end of the inspection season at time T. The paid bonus is based on the revenue τ_j of the contract j from the bonus season and the bonus percentage d_j, which in turn is based on the bonus points for the contract. Bonus points ρ_j are calculated as the average inspection points per inspection, i.e. $\rho_j = x_{w_p,j}(T)/x_{w_c,j}(T)$, where $x_{w_p,j}(T)$ and $x_{w_c,j}(T)$ are the inspection points and inspections at the final state for contract j, respectively. The total incentive bonus is the sum of bonuses from all contracts, denoted as $G = \sum_{j=1}^{J} g_j$, where g_j is the contract specific bonus and can be defined as $g_j = \tau_j d_j$, where the bonus percentage for contract j is calculated as

$$d_j = \begin{cases} 1\% & \text{if } \rho_j \leq l_1 \\ 0.5\% & \text{if } l_1 < \rho_j \leq l_2 \\ 0 & \text{else} \end{cases},$$

where l_1 and l_2 are the lower and upper bonus limits set by the incentive program, respectively. As stated in Section 2.1, the limits are either fixed or undisclosed until the end of the season depending on whether a contract is in the new or old bonus system, respectively.

In the case of our example system, if the new state would in fact be the final state and all contracts had the fixed limits of 10 and 25 bonus points, the bonus percentages for each contract could be calculated as in Table 3.2. Contract $j = 1$ is in the second threshold, and could have gained the best bonus by different actions. Contract $j = 2$ is far from the lower bonus limit of 25 points, which suggests that either the actions done have been insufficient or the contract has suffered from excessive amount of faults. As for contract $j = 3$, it has the best percentage and the largest revenue, yielding a significant portion of the total bonus of 1150 EUR. The total bonus percentage from total revenue and total bonuses is approximately 0.6%.

Table 3.2. The example state would gain a total of 1 150 EUR in incentive bonuses, if the inspection season would end at $t + 1$. Total bonus percentage is then 0.6%.

Contract	Bonus points	Bonus percentage	Revenue	Bonus (EUR)
j=1	12.5	0.5%	30 000	150
j=2	42.1	0%	50 000	0
j=3	7.2	1%	100 000	1 000

3.3 Estimation of parameters of the model

A critical part of simulating the model is to tune the random variables reliably. There are several variables estimated to follow a Poisson point process, more specifically the fault rate per day for each fault category per contract, the inspection rates per weekday per contract, the inspection points from the more severe faults per contract, and the repair rate per day per fault category for each workshop. All of these phenomena have comprehensive historical data to observe the average amount of occurrences per time period.

The public authority, HSL, shares the relevant datasets from inspections and the points gained through the inspections, in addition to data of individual faults indicating how they were handled or repaired. The data used for estimating the parameters for the random variables stretches out two years back to include enough data to draw conclusions, but to exclude irrelevant data from the past, since e.g. maintenance speed changes with investments and process development, nullifying the relevance of the past data. Inspection rates are calculated separately for each contract and weekday since inspectors work only from Monday to Friday. On average, there are 0.6 inspections per contract on Tuesdays and Wednesdays, 0.5 inspections on Mondays and Thursdays, and only 0.3 inspection on Fridays. However, some contracts average to over 1.1 inspections per day, while others remain under 0.2 inspections per day, mainly due to less intensive driving schedules.

New minor faults per day are specific for each fault category and contract. Surprisingly, out of the 17 different fault categories, only three have on an average contract over 0.01 faults per day. These faults appear on average 0.05 times per day per contract. Some contracts gain as high as 0.2 new faults per day, while some gain no faults at all in several categories. The difference on new faults origins from different frequency of departures, traffic conditions, and passenger demographics. Inspection points from the more severe faults are contract specific and average to 8.9 points per day, whereas most contracts have 5 points per day. At most a contract gains 47 points per day. Same reasons for differences between contracts for point accumulation hold as they are for inspection rates and fault rates.

Repairs per day are approximated for the two considered workshops separately and for each fault category. The inspection data from HSL includes the date of inspection and the repair date. Extracting the time taken for a repair action in days, averaging over available data points, and taking the inverse of it yields an average amount of repairs per day. Since the maintenance processes are under constant development, the repair data is first examined from the previous three months, from which the values are calculated. However, there are more gaps on different fault categories, so if no value exists, the scope is extended to two years. Despite this, some fault categories still lack the data to calculate the repair time due to never occurring during the two year period. Such categories receive the average repair time for all available categories. On average, different fault categories receive 0.3 repairs per day in a workshop with the highest value being 0.5 and lowest 0.2 repairs.

The results are highly dependent on the output of the Poisson point processes, which is the reason why the simulations are run several times to exclude randomness. The final state of each MDP from the simulations is averaged to calculate the expected performance of each policy. Increasing the number of simulations improves confidence intervals resulting in a more reliable result for the expected incentive bonuses. However, simulations are time consuming. Scripts that take hours to run decrease the usability of the decision tool. Thus, a compromise is required to select a reasonable amount of iterations. The results in Section 4.3 were extracted as the average outcomes from 100 iterations.

4. Results

This Chapter describes how the model operates with different policies, and how those policies were developed in Section 4.1. The details on how the simulations of the model were done are in Section 4.2. The results of the simulation are introduced in 4.3. Finally, a sensitivity analysis is found in Section 4.4.

4.1 Development of policies

As presented in Section 3.1, the quality incentive program as a system is complex and has many restrictions in terms of modeling the system as a Markov Decision Process (MDP). Thus, many approximations and simplifications lead to the definition of a policy used to determine the actions being reduced to describe only the maintenance order between contracts.

The current way of repairing faults is by chronologically fixing them as they come by, on the condition that there are resources to fix them. The focus is on the more severe faults that are required to be repaired in a deadline. If the repair workshops can repair the minor faults, they will fix them. Based on historical data, an average amount of repairs per day can be calculated to model the events as a Poisson process. The process is used to model how many faults one workshop can fix in a day.

Since the modeled system lacks the data of which fault has been observed at which time, the current policy to repair faults can be approximately modeled to the system by repairing minor faults randomly from contracts. Each workshop handles a set of contracts assigned to them, and each workshop has their own rate of repair actions for each fault type per day. This kind of a random policy is a robust way to repair faults in contracts and treat the contracts equally. However, the policy is not necessarily the best in terms of gaining the most incentive bonuses from all contracts, since contracts are not equal between each other due to accumulating different amounts of revenue. Other possible policies focus differently on contracts, which changes the state of the system and in the end may change the amount of bonus received. Constructing other policies may be in some cases straightforward, and in some cases quite arbitrary. There are vast numbers of possibilities for different policies, so only a small set is formed for an approachable comparison between the policies, and rational interpretations of how those policies are implemented.

A greedy alternative to the current policy of random order is to have the contracts in the order of accumulated revenue. This maintains the largest contracts on the best possible bonus level due to low amount of faults caught in inspections. The trade-off is that the focus could prove to be too much in the top contracts, and the middle sized and smaller contracts risk losing all bonuses, which could in the worst case have a negative total impact.

One possible policy could be ranking the contracts based on how many minor faults each has in total. This emphasizes minimizing the inspection points per inspection for all contracts. This policy also has a similar trade-off to the randomly based policy, since focusing on low-revenue contracts yields lower incentive bonuses compared to the high-revenue contracts. However, lower revenue contracts have less traffic on them, which correlates positively with the amount of minor faults.

Another possible policy is to assign the repair actions to contracts which are on the verge of losing or gaining bonus. The contracts are organized in an ascending order of their relative distance from the nearest bonus threshold limit. The closer a contract is to the closest limit, the higher priority it receives on the maintenance queue. Benefits to this method are that the efforts have likely the most impact due to gaining a bonus with few actions, or not losing bonus due to negligence on any contract. The disadvantage is that the policy disregards contracts that are far from their respective bonus limit. On the one hand, such contracts could be thought as so called lost causes, such as contracts with 50 or more bonus points, but on the other hand, there might be contracts that could give a major proportion of the incentive bonuses if maintained more thoroughly.

Finally, combining the best effects and losing some of the trade-offs of previous policies would be to rank the contracts based on how close they are to gaining or losing bonus, that is the relative distance to a bonus limit, and how much the revenue is for the contracts. Basically, each contract receives a rank score r_j as $r_j = \frac{\tau_j}{\delta_j}$, where $\delta_j = \min(\frac{\rho_j - l_1}{l_1}, \frac{\rho_j - l_2}{l_2})$. Thus, if any lost causes exist based on the rank scores, they are more likely to have low revenue.

4.2 Simulating the model

The model must be simulated in order to extract any results due to the random variables. The simulation is run from the first day of the inspection season from the fall of 2020 to the end of the inspection season with the initial state build from realized data. The goal of the simulation is to calculate the expected total incentive bonus at the end of the inspection season for different policies. Due to the random variables, the simulations are run 100 times with a set of parameters to mitigate variance and improve the confidence intervals. From the set of simulations, the received bonuses are averaged to receive the expected value of incentive bonuses by following any policy. All contracts belonged to the old system of bonus thresholds, and the limits are known retrospectively, which were 10 bonus points for the best third and 22 bonus points for the second best third in the fall of 2020. Thus, the bonus thresholds for each contract can be determined based on the realized threshold limits.

There are several aspects to consider while building the simulation. First, the simulation requires the initial data of what is the situation in the system, which acts as the basis for the simulation. Second, each tested policy has a unique MDP due to different actions on the process, however, the number of inspections, new faults and repair actions are the same between the policies, since a policy does not affect any of them.

The model was built as a Jupyter Notebook, and the simulations were run with Visual Studio Code on a personal laptop, resulting to approximately 75 seconds per simulation. The Poisson processes that each random variable follows were simulated with the Poisson module from Scipy package. The realization of each process was determined by using the cumulative distribution function cdf-method and a uniform random variable uniform-method from Numpy package.

4.3 Performance of policies

The performance of a policy is measured by the accumulated incentive bonus at the end of a bonus season. Table 4.1 shows the comparison between the different policies for the bonus season of fall 2020. The policy with the ratio of revenue and distance to bonus limit offers the largest bonus from the tested set of policies. Compared to the random situation, the improvement is 8.2% with the best policy. The policy with focus on the amount of minor faults offers the lowest bonus. In fact, the worst policy appears to do significantly worse with approximately 8% less bonus compared to simply randomly repair faults.

Table 4.1. Results of the bonus season with different policies. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Model	Random	Revenue	Fault	Limit	R&L
Bonus-%	0.29	0.30	0.27	0.31	0.31
Bonus (kEUR)	156	159	143	166	168
Difference to random (%)	0.0	2.0	-8.0	6.7	8.2

With the results from the sensitivity analysis, it is possible to compare the policies with even more simulations, as any change of the tested parameters should not affect the differences of the policies. Table 4.2 shows the average results from all simulations.

Table 4.2. Average results of incentive bonus from all the simulations including the ones from sensitivity analysis. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Model	Random	Revenue	Fault	Limit	R&L
Difference to random (%)	0.0	1.3	-7.3	8.1	8.9

4.4 Sensitivity analysis of the estimated parameters

The model has several random variables that are modeled with historic data. However, historical events are not a guarantee for future events. The expected realizations of the random variables based on historical data might in fact deviate from the actual average realizations. For example, the Covid-19 pandemic reduced the amount of commuting, which could have lead to fewer faults occurring due to lower passenger counts and fewer traffic. To test the impact of deviances of the estimated parameters, a sensitivity analysis is conducted to demonstrate how the results are affected by each parameter.

The analysis is executed by running a new set of simulations and changing one parameter at a time by 10%, 20%, -10%, or -20%. The examined parameters are the random variables, i.e. inspection rates, fault occurrence rates, repair action rates, and inspection point rates of the more severe faults, which all have several values for each contract or each fault category. The changes are applied to all different values within the parameter. For example, the change of 10% in inspection rate results to more inspections on average with every contract. Each change is run with a unique set of a hundred iterations, which unfortunately means that the realizations of unchanged parameters vary between each change of the investigated parameter. Ideally, all unchanged parameters would be controlled between the changes to observe changes only caused by the change of the investigated parameter.

The results for the analysis are presented in Tables 4.3-4.6. The numbers represent the relative changes of the received incentive bonuses compared to received bonuses with original parameters found in Table 4.1. Each Table includes the average change between the policies for each parameter change.

The inspection rate has arguably the most interesting role of the parameters. On the one hand, with no inspections there would be no inspection points, and on the other hand, inspections with little inspection points decrease the bonus points. Thus, in some cases inspections improve the results, and in other cases they worsen the outcome depending on how many faults a bus has during an inspection. Table 4.3 shows the sensitivity analysis on the inspection rate, which suggests that fewer inspections in fact decrease the received incentive bonus and vice versa. However, the results do vary, and the magnitude of the impact is small compared to the impact of e.g. fault occurrence rate. The difference between the values received from $\pm 20\%$ changes for each policy is relatively constant, approximately 3.3% percentage points. The analysis produces no strong conclusions, however, it suggests that the impact of inspection rate is small.

Table 4.3. Results for the sensitivity analysis on the inspection rate. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Change	Random	Revenue	Fault	Limit	R&L	Average
+20%	1.5%	0.6%	2.3%	2.8%	2.5%	1.9%
+10%	-0.1%	-0.3%	0.1%	1.0%	0.0%	0.1%
-10%	-2.3%	-2.1%	-1.7%	-1.6%	-1.1%	-1.8%
-20%	-1.2%	-2.4%	-1.1%	-1.1%	-1.0%	-1.4%

The amount of possible repair actions is limited, however, the repair processes can be improved. The effect of repair actions appears to be relatively small based on the results on Table 4.4. The results are not consistent on magnitude which shows that the results are susceptible to deviation, however, the more negative change in repair actions, the less bonus is received and vice versa. Improving the repair processes should then pay off, and since the repair rate is a parameter under the operator's control, it is encouraged to focus on it. Based on the analysis, one could argue that the most expensive repair actions could be ignored since the impact of the repair action rate is not that significant, however, this requires further study and model development.

Table 4.4. Results for the sensitivity analysis on the repair action rate. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Change	Random	Revenue	Fault	Limit	R&L	Average
+20%	1.9%	2.1%	0.8%	3.2%	3.5%	2.3 %
+10%	-2.0%	-0.1%	-1.3%	0.0%	-0.3%	-0.7%
-10%	-3.0%	-3.4%	-0.7%	-1.0%	-2.0%	-2.0%
-20%	-2.8%	-4.8%	-1.7%	-4.1%	-5.2%	-3.7%

The rate of faults has a clear impact on the received bonus as is shown in Table 4.5. The more new faults occur, the less bonus is received. The results are quite consistent in terms of the direction and magnitude. Historically, approximately half of the inspection points are gained from minor faults, and the rest from the more severe faults, which fits quite well for the magnitude of the change.

Table 4.5. Results for the sensitivity analysis on the fault occurrence rate. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Change	Random	Revenue	Fault	Limit	R&L	Average
+20%	-7.5%	-10.7%	-6.5%	-7.9%	-9.5%	-8.4%
+10%	-5.5%	-6.0%	-4.9%	-4.6%	-5.7%	-5.3%
-10%	3.0%	3.8%	3.4%	6.9%	6.4%	4.7%
-20%	8.7%	8.2%	9.4%	10.8%	10.6%	9.6%

The sensitivity analysis on the inspection points rate of the more severe faults produces expected results found in Table 4.6. Increasing the rates reduces the received bonus, since more inspection points increases the bonus points, while reducing the rates increases the received bonus. The magnitude is in line with how approximately half of the inspection points from the more severe faults.

Based on the sensitivity analysis, the two most sensitive parameters are the fault occurrence rate and the inspection points rate of the more severe

Table 4.6. Results for the sensitivity analysis on the inspection points rate of the more severe faults. Model names refer to the focus points of each policy. The fifth policy with the ratio of revenue and difference to bonus limit is denoted as R&L.

Change	Random	Revenue	Fault	Limit	R&L	Average
+20%	-7.4%	-12.0%	-6.1%	-4.8%	-6.9%	-7.4%
+10%	-5.0%	-5.9%	-3.5%	-1.8%	-2.6%	-3.8%
-10%	1.5%	2.3%	1.4%	3.5%	2.2%	2.2%
-20%	7.7%	7.3%	8.9%	8.8%	7.9%	8.1%

faults. Both of the parameters are related to faults that are found on an inspection. If the faults could be detected and repaired before caught in an inspection, the impact on the incentive bonus would be large. Once the faults have been caught once, the amount of repairs done appears to have a relatively low impact. As for the inspection rate, the impact of it is fortunately low, since the operator is not able to control the variable.

5. Discussion

The objective of the thesis was to develop a decision support tool for fleet maintenance at Nobina to improve the efficiency of the maintenance order by procuring incremental incentive bonus from the quality incentive system. One part of the objective was to develop a set of different policies that could be executed with relative ease and which one of them is the best. Based on the results, the policy with both revenue and distance to bonus limit as the focus points is the best option. The sensitivity analysis supports the superiority, since the analysis produced more simulations to compare the outcomes. Another part of the objective was to create a tool to determine the maintenance order based on the best policy at any given time. In retrospect, the objective was fulfilled.

However, due the course of the thesis project the model of the system experienced several simplifications that might have decreased the accuracy of the model. For example, the model handles faults at a contract level, although they exist at bus level. There is a relatively large variance on how many faults a bus may have, for example the age of the bus correlates positively with the amount of faults. This directly affects how many inspection points a bus may incur from an inspection. In addition, buses can be switched between contracts with certain strict vehicle requirements. For example, most of the buses can be switched with electric buses due to HSL valuating them high, whereas some bus lines travel in the highways with more capacity, creating a need for large buses. Both necessary and voluntary bus switches occur naturally due to e.g. accidents, maintenance schedules, or a surplus of more efficient vehicles. Implementing this side could prove to be the very beneficial for more reliable and better results as an entire new set of actions is introduced to the model. This would require researching the vehicle requirements thoroughly and creating and maintaining a database for the information.

Another deficiency in the model is the lack of costs for actions. The costs would affect the model by giving a more realistic outcome of the actions, giving the decision makers a better tool. Fault categories have very varied costs, ranging from tens of euros to over a thousand euros. Thus, focusing on faults with a good inspection points per repair cost ratio could give rise to new and better policies. The issue with these costs is that there is no way to reliably estimate a cost for a single fault in the dataset, as there is no additional information on the nature of the fault which would be required to create a clear picture of what ought to be done to repair the fault. It is possible that many faults can be repaired at the cost of one action, e.g. several scratches on a window only require changing the one window. Consequently, the costs have a high variance, which could be harmful to the model overall. If these costs and bus switching would be introduced, the costs for bus switching could also be implemented with relative ease due to more stable and predictable behavior. The costs of bus switching are low, as traveling costs of the buses are cheap compared to the costs of repairing faults. If bus switching would be allowed to incur the sanctions related to it, they would have a more significant impact, but it would provide a more diverse set of feasible actions.

In addition, the number of faults is difficult to know with certainty for several reasons. First, the inspectors may interpret the number of faults differently, e.g. two scratches next to each other may be marked as one or two faults occasionally. Second, some faults are caught multiple times, each time yielding more inspection points. However, the data lacks the necessary information to interpret whether a new fault has in fact been caught in a previous inspection. Thus, each inspection accumulates inspection points, but does not necessarily add new faults to the state. Finally, fixing many faults at one repair action changes the amount of existing state too. It is practically impossible to approximate the frequency of such events with the existing data, and more detailed recording of faults and maintenance is required for such conclusions.

The assumption that inspections are done randomly for buses for each contract should be a good assumption, however, inspectors have access to the inspection system where open faults may be observed, and inspections are possible to focus on buses with many open faults. Similarly, the assumption that all existing faults are caught in an inspection might not hold, since passengers may cover some of the faults, and the inspector might not have access to every part of the bus, such as the left side of the bus from the outside due to traffic.

The time step of one day fits for the purposes of making decisions well, but for example the inspections and emerging faults differ based on the time of the day, which has some effect on the inspection points gained from an inspection. The model treated this spread by assuming half of the new faults in a day can be caught in inspections, whilst the other half occurs after the inspections are done for the day. In addition, weekdays have their differences due to driving schedules and passenger behavior, but this aspect was taken into consideration as weekdays were handled separately.

In the model, many events were approximated as Poisson processes, namely new faults on buses, inspections, inspection points from the more severe faults, and fixes a workshop can perform. In order for the variables to truly follow a Poisson distribution, the events need to be independent from each other. Especially the new faults and fixes might in fact have dependency in some cases, as one fault may lead to another fault, or fixing one fault might require fixing another fault – or it just is convenient to do at the same time. These cases, however, should not have a significant impact based on their scarcity.

In addition, HSL sets a target number of inspections for the inspectors, which means that if inspections are short of the target at the end of the inspection season, there will be more efforts to fulfil the target. This leads to special cases of the distribution, such as the zero-truncated Poisson distribution, where the number of events is a positive integer. This or other special cases could be utilized at the company level, but since the inspections need to be modeled at the very least on contract level, it is impossible to determine which contracts would be affected by the special case. Hence, the inspection target is not considered in the model. In fact, the average number of inspections from simulations falls under the target value set for Nobina. One explanation is that since the target level has risen in the past two years, the historical data produces a lower rate of inspections than what is necessary to produce enough inspections.

The sensitivity analysis had an issue on controlling the random variables that were not examined, as each change in a parameter resulted in a whole new set of simulations with different realizations on the random variables. Technically, it would have been possible to build the simulation script with the ability to control all parameters, but since the sensitivity analysis was conducted at the end of the project with little time to spare, the control was excluded.

6. Conclusions

The objective of this thesis was to develop a decision support tool for a quality incentive program at Nobina to determine how the maintenance order could be improved for additional incentive bonuses, and giving a recommendation on what the order should be at any given situation. Another benefit of the model was to form a process to gather information from various sources to produce a clear vision of the current state at any point, which was previously concealed.

The results suggest that it is possible to improve the current maintenance order of a relatively random way by taking account the size of a contract by revenue, and how close to gaining or losing bonus the contract is. This policy proved to be the best of the five policies tested in thirteen out of fourteen sets of simulations including the simulations from the sensitivity analysis.

The research done in the thesis provides a solid foundation for further development of the processes around the system and improving the model. As it stands, the model is a stripped version of the system with several simplifications. With more precise and thorough datasets, the model could be improved greatly to consider e.g. the costs of the actions to provide a more realistic view of the system. In addition, instead of comparing pre-determined policies, it is possible to conduct policy optimization, which could likely improve the outcome considerably.

Bibliography

Aviv, Y., & Pazgal, A. (2005). A partially observed Markov decision process for dynamic pricing. *Management Science*, *51*(9), 1400–1416.

Bellman, R. (1957). A Markovian decision process. *Journal of Mathematics and Mechanics*, 679–684.

Cha, J. H. (2018). Point processes for reliability analysis : Shocks and repairable systems. *Cham, Switzerland*, Springer.

Date, S. (2019). The Poisson Process: Everything you need to know. Retrieved April 15, 2021, from https://towardsdatascience.com/the-poisson-process-everything-you-need-to-know-322aa0ab9e9a

Dixit, A. K., & Sherrerd, J. J. (1990). *Optimization in economic theory*. Oxford University Press on Demand.

Durrett, R. (2019). *Probability: Theory and examples* (Vol. 49). Cambridge University Press.

European Comission. (2020). Urban Mobility. Retrieved August 7, 2020, from https://ec.europa.eu/transport/themes/urban/urban_mobility_en

Guttorp, P., & Thorarinsdottir, T. L. (2012). What Happened to Discrete Chaos, the Quenouille Process, and the Sharp Markov Property? Some History of Stochastic Point Processes. *International Statistical Review*, *80*(2), 253–268.

HSL. (2020). HSL-alueen asukkaiden tyytyväisyys joukkoliikenteeseen pysyi viime vuoden BEST-tutkimuksessa korkealla tasolla. Retrieved September 25, 2020, from https://www.hsl.fi/uutiset/2020/hsl-alueen-asukkaiden-tyytyvaisyys-joukkoliikenteeseen-pysyi-viime-vuoden-best

Itô, K. (2020). Poisson point processes attached to Markov processes, In *Contributions to probability theory*. University of California Press.

Jansson, K., & Pyddoke, R. (2010). Quality incentives and quality outcomes in procured public transport–Case study Stockholm. *Research in Transportation Economics*, *29*(1), 11–18.

Kristensen, A. R. (1996). Textbook notes of herd management: Dynamic programming and Markov decision processes. *Dina Notat*, (49).

Kuukankorpi, A. (2020). Hsl:n liikennöintisopimukset. Retrieved August 10, 2020, from http://www.kuukankorpi.com/paikallisliikenne/hslsopimukset.html

Leppinen, J. (2020). *A Dynamic Optimization Model for Maintenance Scheduling of a Multi-Component System* (Master's thesis). Aalto University.

Markov, A. A. (1954). The theory of algorithms. *Trudy Matematicheskogo Instituta Imeni VA Steklova*, *42*, 3–375.

Pääkaupunkiseudun Kalusto Oy. (2020). Yritys. Retrieved August 10, 2020, from http://junakalusto.fi/fi/yritys

Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *1988*, *64*(1), 12–40.

Parks, M. (2017). The History of Public Bus Transportation. Gogo Charters. Retrieved August 7, 2020, from https://gogocharters.com/blog/history-of-public-bus-transportation

PwC. (2020). Rapid Urbanisation. Retrieved August 7, 2020, from https://www.pwc.co.uk/issues/megatrends/rapid-urbanisation.html

Song, H., Liu, C.-C., Lawarrée, J., & Dahlgren, R. W. (2000). Optimal electricity supply bidding by Markov decision process. *IEEE Transactions on Power Systems*, *15*(2), 618–624.

Stirzaker, D. (2000). Advice to hedgehogs, or, constants can vary. *The Mathematical Gazette*, *84*(500), 197–210.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Thai, V. V., Tay, W. J., Tan, R., & Lai, A. (2014). Defining service quality in tramp shipping: Conceptual model and empirical evidence. *The Asian Journal of Shipping and Logistics*, *30*(1), 1–29.

The Board of HSL. (2019). Toiminta- ja taloussuunnitelma 2020-2022. Retrieved August 10, 2020, from https://www.hsl.fi/sites/default/files/hsl_-_toiminta-_ja_taloussuunnitelma_2020-2022.pdf

White, D. J. (1993). A survey of applications of Markov decision processes. *Journal of the Operational Research Society*, *44*(11), 1073–1096.