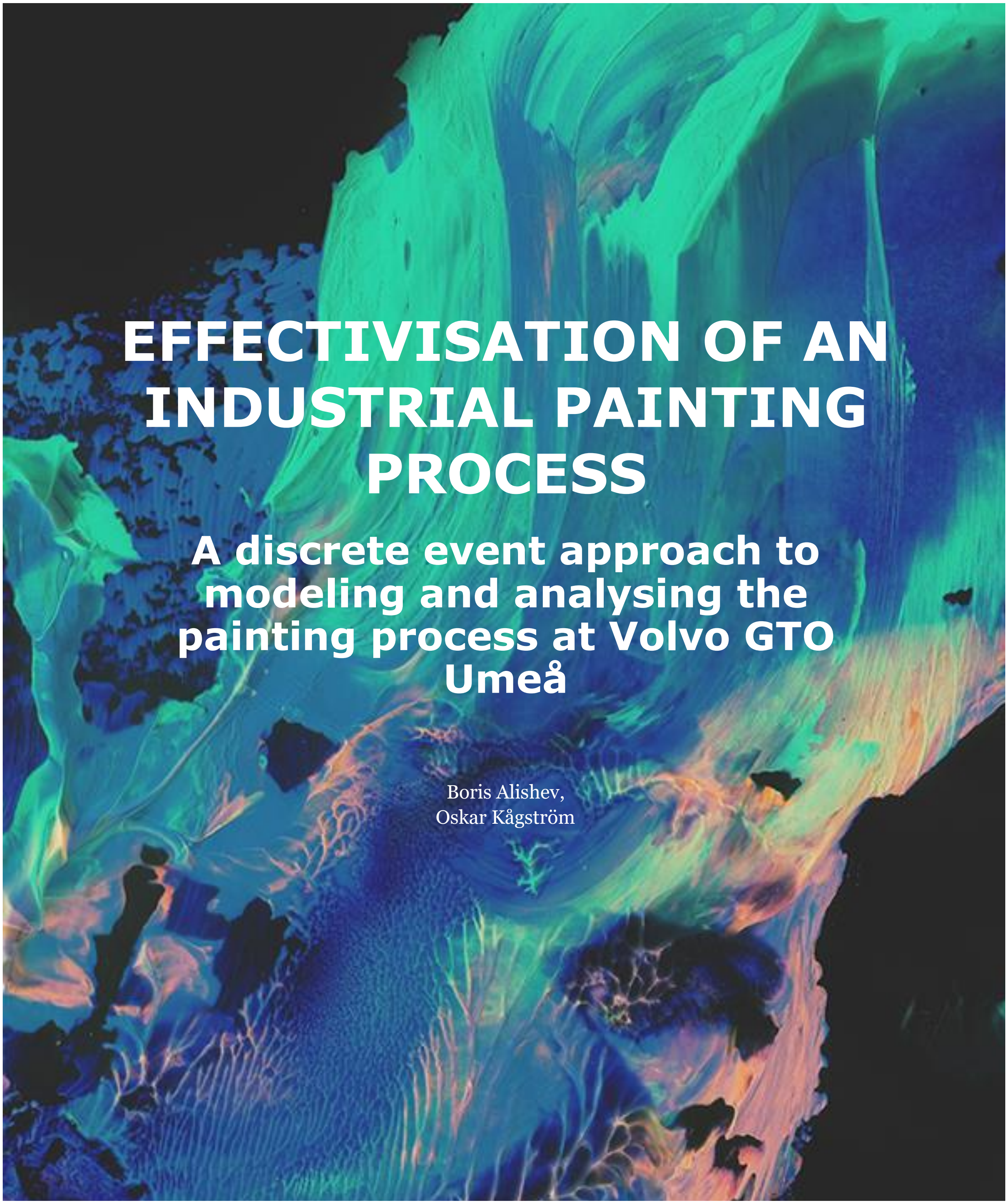
UMEÅ UNIVERSITY



EFFECTIVISATION OF AN INDUSTRIAL PAINTING PROCESS

A discrete event approach to modeling and analysing the painting process at Volvo GTO Umeå

Boris Alishev,
Oskar Kågström

Master thesis, 30 ECTS
Master of Science in Industrial Engineering and Management, 300 ECTS
Spring term 2022

EFFECTIVISATION OF AN INDUSTRIAL PAINTING PROCESS
Master thesis for Volvo GTO & Umeå University

Submitted in partial fulfilment of the requirements for the degree Master of Science in Industrial Engineering and Management
Department of Mathematics and Mathematical Statistics
Umeå University
SE - 907 87 Umeå, Sweden

Supervisors:
Konrad Abramowicz, Umeå University
Torbjörn Hallén, Volvo GTO Umeå
Examiner:
Patrik Rydén, Umeå University

Abstract

For any manufacturing process, one of the key challenges after a solid foundation has been built is how improvements can be made. Management has to consider how possible changes will affect both the process as a whole in addition to every individual part before implementation. The groundwork for this is to have a clear overview of every part and the possibility to investigate effects of changes. This thesis thus aims to provide a clear overview of the complex painting process at Volvo GTO in Umeå and a template for investigating how differently implemented changes will affect the process. The means for doing this is to use statistics, modeling and discrete event simulation. Modeling shall provide an approximate recreation of reality and the subsequent analysis shall take into account similarities and differences to estimate the effects of changes. Recreation of real-world data and variability is based on bootstrap resampling for multiple independent weeks of observations. Results obtained from simulation are compared to observed data in order to validate the model and investigate discrepancies. Given the results of model validation, modifications are implemented and information obtained from model validation is used to evaluate the results of the modifications. Further, strengths and weaknesses of the thesis are presented and a recommendation of altering the stance on process improvements is provided to Volvo GTO.

Keywords— Bootstrap, Resampling, Descrete Event System, SimEvents

Sammanfattning

En av de viktigaste utmaningarna för industriföretag är att tillämpa förbättringar till redan fungerande processer. Ledningen måste där ta hänsyn till hur möjliga förändringar påverkar inte bara processen som helhet, men också hur varje enskild del påverkas. Grunden för detta är att ha en tydlig överblick över varje delprocess samt möjlighet att undersöka vilka effekter som förekommer av olika förändringar. Detta examensarbete syftar till att ge en tydlig översikt över den komplexa målningsprocessen på Volvo GTO i Umeå och ett medel för att undersöka hur olika genomförda förändringar kan komma att påverka processen på olika vis. Medlen för att göra detta är att använda statistik, modellering och diskret händelsestyrd simulering. Modelleringen ska ge en ungefärlig skildring av verkligheten och den efterföljande analysen ska ta hänsyn till likheter och skillnader för att uppskatta effekterna av implementerade förändringar. Återskapande av verkliga data och variation baseras på en återsamplingsmetod (bootstrap resampling), applicerad på flertalet oberoende veckor av observationer. Resultat som erhållits från simulering jämförs med empirisk data för att validera modellen och undersöka avvikelser från verkligheten. Då resultaten från simuleringar validerar modellrepresentationen testas möjliga modifikationer i modellen där information från modellvalideringen används för jämförelse av modellerna. Vidare presenteras arbetets styrkor och svagheter och en rekommendation om att ändra synen till hur processförbättringar bör genomföras ges till Volvo GTO i Umeå.

Svensk titel: Analys och modellering av en industriell målningsprocess

Acknowledgements

We wish to thank various people that have helped contribute to this thesis. Firstly, we would like to thank our supervisor Torbjörn Hallén. You have given us continuous support and helped us feel at home at the facility and we will cherish the insight that you shared.

Extreme gratitude goes to our supervisor Konrad Abramowicz for his enthusiasm, knowledge, patience and time, both during weekdays and weekends. You have made us work tirelessly to complete this thesis even when the finish line seemed nowhere to be found, and we could not be more grateful for it. Your support has been greatly appreciated and we have thoroughly enjoyed the conversations we have had with you.

We would also like to thank our families and friends for their support and understanding as we may have seemed distant while working on this thesis. Special appreciation goes to our classmates who have made the studies during the past five years fun and enjoyable.

Finally, we would like to express our gratitude towards all employees at Volvo GTO that have been helpful in answering our questions, despite how repetitive they may have been.

Contents

1 Introduction

1.1 Background

In an era where customer demands for product quality is continuously increasing, bigger stress is placed on manufacturers to increase production speed, while also improving the quality of the products. For industries with high costs per entity, this is particularly important as the cost of mistakes cannot be overshadowed by sales volume. For Volvo GTO, this means that high emphasis has to be placed on both having an automatic, standardised process, as well as following up on said process to ensure that things have gone according to plan and if not, alter the issues in the most cost and time-effective manner possible. In a perfect scenario, all employees involved in these quality ensuring processes would be of similar expertise and produce uniform results. This is however very difficult to accomplish. Due to personnel reforms constantly occurring, the level of expertise in different areas vary significantly. As such, decisions taken by one employee based on a specific issue can be completely different from what another employee would decide upon. The result of this is that additional errors other than the mechanical ones occur from the processes implemented to correct said errors. As these errors and decisions are not black and white, straight up implementing a top to bottom curriculum is essentially impossible. Therefore, Volvo GTO in Umeå are looking into whether an in-depth overview of the painting system may allow for a more precise framework for specific actions to be implemented. As faulty actions can lead to massive costs, a way to simulate the effects of these implementations is desired to ensure that they do not provide negative results.

1.1.1 Painting System Description

The flow at the paint shop starts from cabs arriving from a previous building where the cabs are constructed and prepared for the painting process. The cab transports through the paint shop on a skid together with its specific plastic parts which in the end are going to be assembled together with the cab before being delivered to the customer.

In the first process the cab (together with its plastic parts) is brushed in order to remove any possible dirt or dust before advancing into the primer painting process. There are two lines for the primer painting which then lead to each of the two respective ovens for the paint to dry. Ovens are located on a level above (12-meter level) to where the cabs are transported by each respective elevator. The ovens have the maximum of 12 places each for the cabs, and after that the cabs are transported back down by elevators to the original level (7-meters level) to the control stations. Previously these control stations were used to check the quality of every cab after the primer painting process, but they are not active anymore. However, some cabs are still checked and examined here manually in a case of any suspicion of deviations resulted from earlier processes or poor quality of primer painting.

After that the cabs arrive to an elevator which handles the transport between three levels of the paint shop - 0 meters, 7 meters and 12 meters. Usually the cabs advance past the elevator to continue on 7 meters level forward to "COB" - a color-optimization buffer where the cabs are stored and sorted prior to the next process.

However, if any deviations were discovered after the primer paint process then the cab is transported down to 0-meter level. Two adjustment stations, used in order to repair small deviations, and one truck table is located here. The truck table is used as a waiting spot for a cab when only a part of it needs to go through a painting process again. If the deviations are successfully removed then the cab is transported back up to the 7-meter level to then advance forward to the COB. However, if there is a deviation in primer paint, the cab is required to go through the primer processes again. There is also a possibility that the discovered deviations are not repairable, and in that case the cab is removed from the system to be scrapped.

Another alternative path from the elevator after the primer paint processes is 12-meter level. It consists of a primer coat buffer to which the cabs arrive if, in the case of cabs without deviations, there is no room in the COB and the transport line before it, or, in the case of cabs having a deviation, if there is no room at the 0-meter level. These cabs are stored in the buffer until there is enough room in the systems ahead to be transported there.

In the COB the cabs are sorted into sequences based on their intended color before they advance forward. Thus, the batches of 3 to 5 cabs of the same intended color are created here, in order to optimize the process by lowering the frequency of changing paint of different colors in the top coat processes. From the COB the cabs advance forward to the top coat painting process.

Here the cabs go through the similar processes to the primer paint system. However, when it comes to painting lines there are two identical top coat lines and one base coat line. The base coat line is used for the cabs that are ordered to have metallic top coat. These cabs go through base color process before being transported back to one of the two top coat lines to get a clear lacquer on top of its base coat. This way the metallic cabs proceed through the painting process here twice, while the rest of the cabs only go through one on of the top coat lines before moving forward to the next process. After going through the cooling zone, located after each oven, the cabs arrive at the quality control station.

Here the cab is manually checked for any possible deviations, and if any deviations are detected, they are registered manually to the system and the cab is transported by an elevator down to 0-meter level where these deviations are meant to be fixed. Another possible destination from this elevator is up to the top coat buffer at 12-meter level. Cabs arrive here in the case of the cab having no deviations and also in the case of 0-meter level, which serves as the cabs destination if it possesses any deviation, being full. The cabs are stored here at the buffer before proceeding down to the next processes. If the does not carry any deviations, it is then transported back down to 7-meter level to the final transport line to the next building. Here the flow through the paint shop for the cabs is finished.

However, the path for a cab becomes more complex if any deviations have been located and it gets transported down to the 0-meter level after quality control. The exact destination here on 0-meter level depends on the deviation:

- Adjustment stations - 3 such stations, cabs arrive here in the case of small deviations, often the cabs are polished here. If the deviation shows to be too difficult to adjust here, the cab can be sent to improvement or grinding processes for further repair.
- Improvement stations - 3 such stations, cabs arrive here in the case of bigger deviations, for example if some part of the cab needs to be repainted.
- Panel improvement station - cabs arrive here also in the case of bigger deviations, but only if it is the cab's panel that requires the improvement.
- Grinding station - cabs arrive here due to some large deviations, often if there is a large part of the cab that needs to be polished. After undergoing a successful grinding process, the cab is directed back to COB to then be repainted. If the grinding process did not succeed in solving the problem, the cab is deemed to be incorrigible and is removed for scrapping.

- Masking station - cabs go through this station to be prepared for the improvement station processes, but most often for the panel improvement station for which some parts of the cabs need to be covered.

When the cab is fully adjusted and improved, and does not need to be repainted, it is transported by the elevator up to the 12-meter level to then advance forward to the next building. If the cab is required to be repainted, it is transported back to the COB to go through the top-coat painting process again.

1.1.2 Important Notes About the Process

At Volvo GTO in Umeå, three different types of cabs are produced based on customer orders. Of these three, two are standard in production while the last one is designed based on specific customer orders. All of the models are produced in different variants, however, approximately 98% of the production volume consists of three different variants of the two standard models, FM-model and FH-model, as of today. While the different models of cabs are essentially standardised, the specific color that they are to have is not. There are currently more than 1000 different sets of colors, all while constantly adding new ones and at times, removing old ones. The colors do not only differ over the color spectrum, but also in which specific route through the process they are to take. In particular, colors that require clear coat have to go through a separate painting process to the regular topcoat process before going back in through the ordinary route. The main part of the process is designed to be almost entirely automatic, with only quality controls and thereafter manual adjustments being dependent on human interactions. This is only true in theory though, as a multitude of employees are employed with the purpose of manually altering the flow of the process by stopping and allowing specific cabs to pass through different parts of the process. This manual interference then results in large variance regarding the time each cab spends in each part of the process. The working hours in the factory are divided into different working shifts where different segments have separate break periods. This in turn means that for specific periods during the day, different processes may come to a halt, requiring manual interference from process control in order for the factory as a whole to proceed normally.

1.2 Goal, Purpose and Possible Limitations

Goal of the thesis work is to contribute to a better understanding of the flow in the Volvo GTO's paint shop and an easier identification of the problematic areas and processes which have the biggest impact on the flow and time spend there. This project will result in a model for simulating the flow of cabs in the paint shop where it will be possible to adjust different parameters and see their effect on the time spent in different processes of the flow. Both data analysis, which is used as a foundation for the model, and the model itself create a detailed overview of the painting system and a clarification of the parameters and processes which impact that system. Thus, this thesis work creates a foundation for Volvo GTO to achieve a more effective flow and more standardised processes which can benefit the organisation.

1.3 Motivation for Chosen methods

The decision to build a simulation model came from the agreement with the supervisor from Volvo GTO that a model for simulation of the paint shop would be of great help to oversee the system and understand the effects that different parameters within the system have. The choice of using the MATLAB (2022) toolbox SIMULINK with SIMEVENTS library as a graphical programming environment for the model came naturally from significant previous experience with the program from previous courses and projects during the studies. The concept of using SIMULINK for modeling in industrial processes is also not a new phenomenon as can be seen in Peacock (2002) and Asbjörnsson et al. (2013). Further, the concept that is used for validating

the model, i.e., calculating mean and median values for different parts of the paint shop as randomness has been introduced, is commonplace when modeling real world systems. (Sargent, 2011)
Data analysis was performed with two main purposes: better understanding of the parts in the paint shop, and a crucial statistical foundation for creating the SIMULINK model. The programming language R Core Team (2021) was used for the data analysis, as it is is well-adapted for work with large data frames, provides a large set of statistical tools and overall performs well when it comes to data analysis.

Finally, the practice of introducing randomness in modelling a complex process and using SIMEVENTS as the tool for creating said models, is something that has been proven useful in analysis of business and production environments in a previous thesis, (Björch and Strålberg, 2016), which provides support for the intended method in producing satisfactory results.

1.4 Outline of the Report and Advice for the Reader

The content of this thesis is organised as following. Chapter 2 introduces underlying theory for the methods used. Chapter 3 presents a step by step walkthrough of how data has been processed, how the simulation model has been built and validated against reality as well as how modifications have been implemented in it. Chapter 4 then follows with the resulting numbers obtained from the methodology along with a discussion surrounding their connection to reality. Finally, conclusions based on the thesis and recommendations for Volvo GTO to implement are presented in chapter 5.

2 Theory

This chapter introduces statistical theory and simulation fundamentals that are vital for understanding the methods and results in this thesis, given some prior statistical knowledge. Furthermore, important elements about the software used in this thesis is presented as well.

2.1 Stochastic Simulation

The purpose of a simulation is to study how a system evolves over time. A system can be defined as a collection of entities interacting with each other towards some logical end. The collection of entities varies for different systems depending on what the purpose of the study is. The state of the system can then be defined as the collection of entities needed to describe the system at a specific timestamp. Systems can be categorised into two types, continuous and discrete. In continuous systems the state changes continuously over time with infinitely small time steps. An example of this is how the velocity of an airplane while in the air changes over time. For a discrete system, the system state changes instantaneously at different points in time. An example of a discrete system is a bank where the number of customers inside the bank only changes when someone departs or enters the bank (Law, 2015, p. 3). For a probabilistic system, the studying approach is called stochastic simulation and involves generating and studying the stochastic properties of the system and observing how it evolves over time (Ross, 2013, p.111).

2.1.1 Pseudorandom Numbers

The base for simulation studies is the ability to generate random numbers, taking any value between 0 and 1. While generating truly random numbers is a time consuming task, needed to be done mechanically or by hand, the modern approach of using a computer is to generate pseudorandom numbers. Pseudorandom numbers consist of a sequence of values that are deterministically generated but have the appearance of independent uniform random variables between 0 and 1. The most common method for generating pseudorandom numbers is to start with an initial value obtained from a seed and then recursively calculate successive values to obtain the full sequence (Ross, 2013, p. 39). Thus, due to the mechanics of pseudorandom numbers, changing the seed for each sequence allows for obtaining different random numbers for each simulation.

2.1.2 Bootstrap

In order to utilise data for simulation approaches, knowledge about the correct data distribution and underlying random model is required. While the true random model is not possible to acquire in reality, an approximation can be recreated by using a sampling method, bootstrap. The method is formally explained by Davison and Hinkley (1997, p. 11) by letting $y_1, ..., y_n$ be a single, homogeneous sample of data. "The sample values are thought of as the outcomes of independent and identically distributed random variables $Y_1, ..., Y_n$ whose probability density function (PDF) and cumulative distribution function (CDF) we shall denote by f and F, respectively". When considering using bootstrap methods, there are two types to distinguish, parametric and nonparametric. For a mathematical model with variable parameters or constraints ψ that fully determine f, the model is called parametric and the method based on this is called parametric bootstrap. If no such mathematical model is used, the statistical analysis is nonparametric and is only based on the random variables Y_j being independent and randomly distributed. For nonparametric analysis, an important role is placed on the empirical distribution which sets equal probability n^{-1} for each sample value y_j. The estimate of F corresponding to this is the empirical distribution function (EDF) $\hat{F}$, which can be defined as the sample proportion

$$\hat{F}(y) = \frac{\#\{y_j \leq y\}}{n}.$$

Where $\#\{X\}$ is defined as the number of times event X occurs. Formally the EDF is defined as

$$\hat{F}(y) = \frac{1}{n}\sum_{j=1}^{n} H(y - y_j),$$

where $H(u)$ is the unit step function that goes from 0 to 1 at time $u = 0$ (Davison and Hinkley, 1997, p. 12).

So, for values $y_1, ..., y_n$ that can be assumed to be independent and identically distributed according to an unknown distribution function F, the EDF can be used to estimate the unknown CDF F. As the EDF places equal probability for all original values $y_1, ..., y_n$, each value Y obtained from the EDF is sampled independently from the original data. Thus, the simulated sample $Y_1, ..., Y_n$ becomes a random sample taken with replacement from the original dataset and so sampling from the EDF is equivalent to sampling with replacement from the original data (Davison and Hinkley, 1997, p. 22).

2.1.3 Bootstrap Resampling

The concept of using bootstrap to approximate a population can be summed up by the following in Chihara and Hesterberg (2011, p. 100-102). With a sample of set size n, pull a resample of the same size from that sample, with replacement and then compute a statistic that describes the sample. Repeat the process multiple times and then inspect how the values of the statistic have varied over each resample.
The methodology can thus be used to approximate a larger process by resampling different parts of the process with replacement and then calculate the estimated effect it has on the process as a whole, multiple times.

2.2 Discrete Event Simulation

An approach to simulation brought up in Ross (2013, p. 111-112) is called *Discrete event simulation* or DES. The approach is designed to help simplify following the model over time and to determine quantities that are of interest. The concept of a DES is to utilise variables and events to perform the simulation. In general, the basic model contains three different variables - Time variable, System state variable and Counter variables. The Time variable (t) refers to the time elapsed inside the simulation. The System state variable describes what state the system is in at the time t. Counter variables measure the amount of times that specific events have happened at time t. The DES then functions as such that all of the above mentioned variables update whenever an event occurs in the model and relevant data is collected. There also exists an event list which lists which events will occur in the nearest future and at which time points. As the variables mentioned above update whenever one of these events occur, this allows for tracking how the system evolves over time without looking at all the time points during the simulation.

2.3 Warm-Up Period

Discrete event simulations can be categorised into terminating and non-terminating simulations. For terminating simulations there is a natural event that defines the end of the simulation, thus allowing for easy analysis of the output as there should be no interference from previous simulations. For non-terminating simulations, the situation is not as straight-forward. As the aim of these simulations is to determine the steady-state behaviour of a system, there is no natural event that signals the end of the simulation. Because of this, the output analysis is dependent on the period before the system reaches a steady state and in order to remove that dependency, a warm-up period is required. The warm-up period functions as such that the non-terminating solution starts at time zero but data gathering only begins ones the system has reached a steady state. Thus, removing the output dependency of the warm-up period (Banks and Sokolowski, 2010, p. 54-44)

2.4 Software

2.4.1 Software Usage

During this thesis multiple software programs are used. Microsoft Excel is used to save and transfer data as it is the standard at Volvo GTO. The programming language R is used for processing and analysis of the original data, along with the basis for model validation. For the simulation script and model build, the main software used is MATLAB with the toolboxes Statistics and Machine Learning Toolbox, Parallel Computing Toolbox and Simulink Toolbox with SimEvents library. MATLAB has a multitude of statistical features and features for creating discrete event simulation models.

2.4.2 SimEvents

SimEvents is an additional MATLAB product originating from the Simulink Toolbox. SimEvents is a modelling tool for creating and analysing event-driven processes with its discrete-event simulation engine and component library. SimEvents contains blocks with built-in functions which allows the user to create a system replicating processes in real life with visuals showing the paths objects can take. There is also a system in place that allows the user to collect system describing signals continuously which allows for continuous model analysis. As SimEvents is a product of MATLAB, most of the functionality is also available there which provides the possibility of creating custom functions to pair with the blocks.

Entity, Attribute and Signal

Entity is the object that goes through the system and has a set of *attributes*, which can be defined as *entity's* characteristics. *Attributes* have some initial value at the start which then can be changed during the *entity's* path through the system. *Signal* describes the system's state and status - can be used to analyse the behavior of the system.

The most common and influential blocks used in the model are presented in the Figure 1 and short explanations for them is written below.

Entity Generator

Generates entities, usually at the start of the model, where generation method can either be Time-based (entities are generated based on a specified time-interval) or Event-based (entities are generated when an event occurs).

Scope

Displays the signal that describes the state of the system - it can be the amount of departed/arrived entities, current number of entities in the system etc.

Entity Server

Entities spend a specified amount of time here representing the process where some operations are performed on the entity before it advances further.

Subsystem

Subsystem is a set of other blocks grouped together. It can be created while dealing with larger and complex systems to simplify the structure of the model and to improve readability.

Entity Queue

A place where the entity is stored for a period of time before moving forward. Number of places in a *Queue* in which order entities leave the system can be specified

Conveyor System

Transports entities forward - length of the conveyor, the number of spaces and speed at which the entities move through the conveyor can be specified.

Input Switch and Output Switch

Input Switch takes in multiple number of entities and dispatches them in a single sequence. The order and priority of entering entities can be specified. *Output Switch* takes in entities and directs them in any of the multiple destinations. Number of paths and how that path is chosen can be specified.

Entity Gate

Blocks entities from advancing further depending on a specified criteria. *Gate* is usually activated by receiving a message from some other system which informs if it should be opened or closed.

Digital Clock

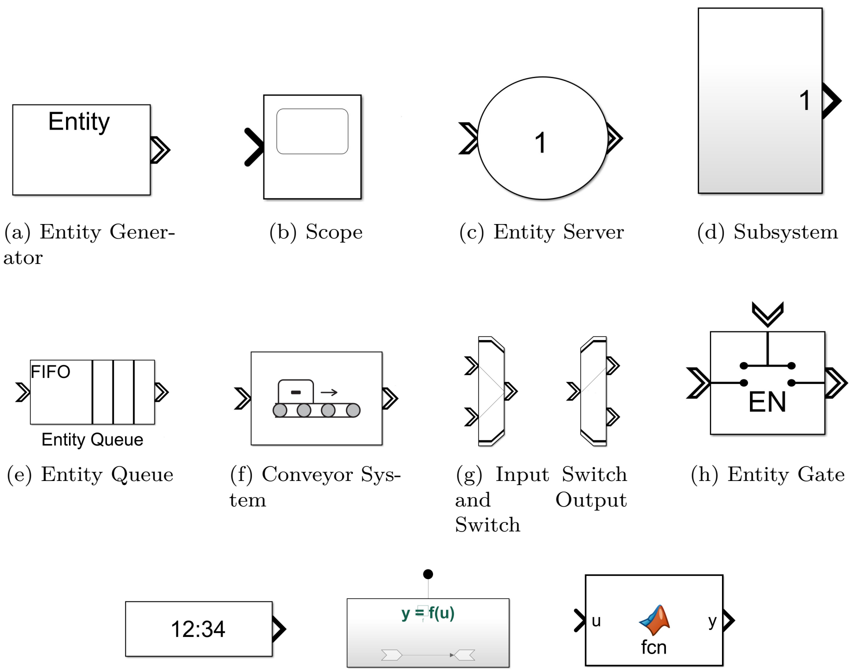
Outputs the simulation time at a specified interval, can be used to calculate the time in the model or parts of the model.

Simulink Function

Creates a function that operates on signal. Often used in a combination with Matlab function which provides an opportunity to implement Matlab code in the model.

Matlab Function

Creates a MATLAB function where input and output signals can be chosen. This allows user to influence the signal in the chosen process and adjust it.



(a) Entity Generator (b) Scope (c) Entity Server (d) Subsystem

(e) Entity Queue (f) Conveyor System (g) Input Switch and Output Switch (h) Entity Gate

(i) Digital Clock (j) Simulink Function (k) Matlab Function

Figure 1: Graphical representation of the SimEvents building blocks used in the model

3 Methodology

This section contains information regarding what is represented in the data, how it has been explored and interpreted along with motivation for the different steps above.

3.1 Understanding the Process

A lot of initial work was invested into understanding the processes and the structure of the paint shop and its complexity. At first a guide through the whole paint shop was given to provide the visual understanding of the different stations and processes, how they are connected and what importance they have on the quality of the produced cabs.

Then, access to Volvo's internal system was provided where it was possible to obtain certain datasets (*QFS* and data about the different colors - both discussed in Section 3.2), but also provided access to the visual structure of whole building and the processes inside of it. This system showed the positions of the cabs in different processes in real time, and thus provided a great opportunity to observe the cabs' path through the paint shop, and contributed with an opportunity to note how these positions are then logged into the system. This connection between the visual system of the processes and the logged data facilitated our work with the data analysis afterwards as it was clearer what the different points in the log data represented.

Afterwards a sketch of the processes in the paint shop was created. It shows a more direct representation of the processes and different paths that a cab can take. The main purposes of this sketch are that it provides an easier understanding of the flow and also serves as a foundation for the *Simulink* model. During the sketch-creation process the idea of a future model was constantly kept in mind so that the model could be easily created having a similar structure.

As the processes were quite complex and thoroughly understanding all possible destinations that cabs could take during their path through the paint shop was critical, regular meetings with process controllers were arranged. These conversations provided a clearer understanding of the details and the important areas and parameters we should pay attention to.

3.2 Data Exploration

The data used in the project is composed of a few different datasets. This is as the system at Volvo records some data for a specific dataset and an alternate amount for another. Apart from this, a dataset containing information about which color number corresponds to what actual color is also necessary to classify different colors. From the different datasets that are automatically recorded, only two are of interest to the project, namely *QFS* and *Event-data*. *QFS* stands for *Quality Follow-up System* and contains data about every deviation for every cab going through the process along with information specific to that cab such as which color (in coded number) and which model the cab is. The *Event-data* records information about when every specific cab arrives at any stage of the process. Important to note is that not all locations are actual processes, instead most of them correspond to different parts of transportation points.

While the data contains every single timestamp of the process, it lacks information about cab model, color and whether or not a deviation has been located, which is why *QFS*-data is vital. For the purpose of explaining as much variation in the process as possible, the main dataset used for analysis was a merged set between *QFS* and the regular log-data. This is to include information regarding amount of deviations, where the deviation occurred, product model and color in order to check for dependence in different processes. The merged dataset was additionally complemented with explanations for the different colors, where designation, clear coat information and amount of laps were the information of interest.

In order to conduct analysis on time spent in processes, the data was ordered by production number and timestamp, in ascending order. In turn, another column was created containing time difference between every logged position, with the first position being set to 0 as that is the start of the process. As data is collected continuously, products can start anywhere in the process which provides an issue with consistency. The reason for this is that the start of the process is labeled as *In primerline*, however, products have the possibility of returning to this point if they are sent back from repairs. Because of this, products that do not originate from *In primerline* were removed from the data in order to simplify the calculations and as only a small number of products are affected, it has little to no effect on the sample size. The process was repeated for the deviation-free dataset.

The data is composed off of multiple different weekly datasets and as such, it is important to ensure that there is no significant difference between the different weeks. A sample of four weeks from different months, not including the period around Christmas, was selected and general analysis of every part of the process was performed. The results were then compared between each week by producing histogram plots and using the *summary* function in *R*. The percentage of deviations was also compared in order to ensure that quality is consistent in the factory. After all of the above measures had been performed and homogeneity between weeks had been confirmed, additional weeks were then selected in the same manner in order to obtain a large sample size. Because the aim of the model is to accurately simulate the regular workflow of the factory, another decision was made to improve homogeneity, namely only including the six standard varieties of cabs as those combine for approximately 98 % of the production volume.

3.3 Analysis

3.3.1 Calculation of Times

Having ordered the data and implementing a variable containing difference in times, a method for calculating each process needs to be created. However, with every location on the path being logged despite if it were part of one singular process, the decision was made to create two algorithms in *R* for calculating times. One algorithm to find the time it took for a process with one single logged station and another for calculating the combined time in a process. The first algorithm named *find_times* was formed through a function which takes an input containing a specific location and returning the first occasion of each cab visiting that location, along with the time difference number and all attributes of the cab. For the second algorithm *find_difference*, there are two functions necessary, one called *find_indices* and the other being the main function *find_difference*. *find_indices* works similarly to *find_times* as it takes a location input and outputs the information connected to the first occurrence of said location for each cab. It does however differ in its output as instead of returning the time difference, it returns the index of the locations. The main function *find_difference* then takes the index output obtained through *find_indices* for the starting and the stopping point of a process as input, calculating the difference between these and returning a vector containing the calculated time and the attributes related to each cab. As the cabs in the starting point do not always match up with the stopping point, the function *match* is used to find the matching cab numbers and those not contained there are removed. Because of the issue where there could be occurrences of a number in the smaller index vector that were not in the bigger one, an additional check had to be implemented. The control uses the *is.empty* function from the rapport package (Blagotić and Daróczi, 2022) in combination with the *is.na* function in order to find missing values and then removing them from the combined vector.

3.3.2 Analysis of Data

After having gathered data about the time each different process takes, further steps needed to be taken before the data could be implemented in the final model. Now, it is necessary to understand that even aside from human interaction, the factory is not homogeneous. Some processes function as a conveyor system while others act individually where another cab cannot begin the process until the one ahead of it has been fully serviced. In addition to this, some processes vary more drastically than others, both due to additional human interactions and due to specific factors regarding the cab. So, the decision was made to firstly investigate which processes were affected by the differences contained in each cab. In order to investigate this, it was required to distinguish the differences between each cab and as we were only looking into standardised cabs, the factors could be broken down into cab variety, color and whether a deviation had occurred or not. While the variety of cab and if a deviation had occurred or not was easily analysed, as there are only six different types and the datasets were easily separated, color is a completely different issue. With more than 1000 different colors being used, obtaining a large enough sample size for all colors in order to conduct an accurate analysis was deemed impossible. Instead, through the aid of supervisors and factory workers, a set of groupings could be determined. The colors were then grouped as those that are most common and easily handled, the colors requiring an additional base coat and the remaining, less common colors.

As the different factors had been determined, the analysis was conducted by using ANOVA in *R*, both with and without interaction between the different factors. While conducting this analysis, the issue of an increasingly lessened sample size became prevalent as equivalent processes are split throughout the factory. In turn, the natural variation caused by stoppages places further importance on large sample sizes. A deep understanding of the factory was therefore important to avoid issues such as there seeming to be a significance regarding color for the first painting process, however, said process takes place before the actual color has been applied to the cab so there was no actual causation between color and time in that process.

Having conducted the analysis of the time spent in every process, the next step necessary to investigate is the proportions of how the data is represented and if there are any dependencies among these. Some of these proportions are: the amount for different cab types, the amount belonging to different color groups, how many of the cabs that obtain deviations during the process, where deviations occur and where cabs are being sent to, depending on where they are being handled. Following the analysis of these, the aim was to create individual cdfs for each individual proportion by combining the percentages for every possible outcome. But, while proportions were easily separated in the higher, more common scenarios such as for cab generation, as the cab progresses further down into the process chain, the sample size drastically decreases. As such, an additional method was used to compare proportions against each other, namely two proportion Z-test.

So, for the processes at the 0-meter section, proportions were compared against each other to investigate whether it would be statistically sound to combine different colour groups in order to obtain a cdf based on a larger sample. However, due to the nature of cdfs, the whole distribution gets disrupted if one part of it is altered. So, the decision was made to only combine samples if all options were deemed not to be statistically different at a significance level 5 %. Another matter to investigate was that cabs have the possibility of going back and forth between different repair sections due to both new deviations occurring as well as cabs simply being sent to the wrong process. The issue at hand here is whether the path probabilities change when visiting a section for the second time, or not. This was again tested with two proportion Z-test to compare the path cabs take directly after attending a repair section to which paths they take after attending that section again, having been at a different section previously.

After analysing the dependencies between these proportions, the general analysis is completed. However, there still remains the decision on how to represent all information in the final model. As has been previously mentioned, the proportions are represented through individual cdfs in different processes of the model. For processes that function as conveyor systems, the estimated standard time for a cab to traverse it without interaction further up the line was picked. For those functioning as elevators or singular carriages, the same methodology was used along with a set percentage of added time depending on how quickly another cab is waiting for that transportation. This is in order to represent the time it takes for the transport to travel back to the starting point after having delivered the original cab. Lastly, for the naturally varying parts of the process, there are mainly two methods for representing the data, finding a distribution or re-sampling. Having attempted to find a theoretical distribution fitting the data, to no success, the decision was made to utilise the large amount of data available and instead re-sample it using bootstrap methodology explained in Section 2.1.2.

3.4 Simulink Model

As it was mentioned previously in Section 2.4.1, program SIMULINK was used to create a simulation model for the paint flow system. In this section the process of creating SIMULINK model is described together with the explanation of how the different parts of the painting processes are represented in the model. The methodology for introducing randomness and dependency to the real state data follows Section 2.1.3 where varying processes are resampled and test statistics for time spent in the model is calculated for each simulation. A layout of the model's most important parts is shown in Figure 2.

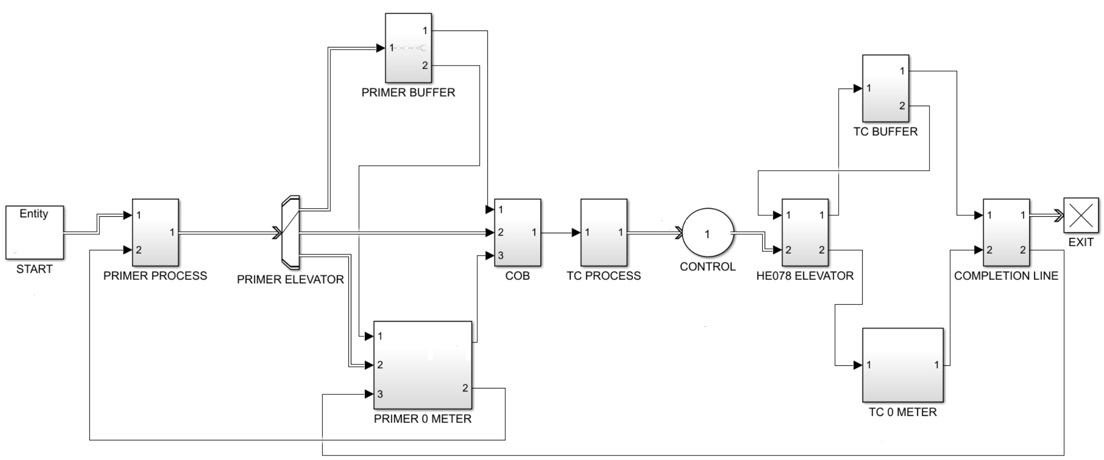


Figure 2: *A layout of the model in* SIMULINK

3.4.1 Cab Generation in the Model

The model starts with an *Entity Generator* producing the *entities* which represent the cabs. The generation method of *Entity Generator* is Time-based which means that new entities are generated after a specific time period in simulation. Said time period is provided by bootstrapping values with replacement from a vector of empirical data. The vector contains observed differences in time from when cabs enter the system chronologically.

3.4.2 Attributes

All *entities* are generated with 17 different attributes: 2 of them represent a specific quality of the cab just like it is in reality, 9 other attributes are implemented with a purpose to guide or steer the cab through specific processes in the model depending on the value of these attributes or in most cases a combination of these attributes' values. Finally, the last 6 attributes are used to measure the time in different parts of the model in order to assess the plausibility of the time compared to the actual observed time in the different processes of the Volvo paint shop. All attributes are shown in Table 1.

Table 1: Attributes used in the model

Attribute Name	**Description**
Type	Model of the cab, 1 to 6
Color	Color group, 1 to 3
Error	Indicates if a cab has a deviation or not
Retur	Indicates if a cab has to go back to the top coat 0-meter level after adjustment station
Skrot	Indicates if a cab has to be scrapped
Prio	Indicates a priority at the Truck table, if a cab should go through or be stored there
Special	Indicates if a cab should go to grinding, masking or improvement station
Insp	Indicates if a cab has to go to an Adjustment station or not
Panel	Indicates if a cab should go to panel improvement station or not
Slipad	Indicates if a cab that has been through the grinding station has gained more deviations or not
Varv	Indicates if a cab is routed to go through the COB a second time
Time	Measures total time of the model
Time_efter_GL	Measures the time between the start of the model and primer elevator
Time_TL_EMU	Measures the time between the primer elevator and the start of top coat process
Time_Kontroll	Measures the time between the start of the top coat process and the control station
Time_TT300	Measures the time between the control Station and the start of the completion line
Time_OK	Measures the time between the start of the final completion line and the end of the model

Attributes **Color** and **Type** are the ones that represent real characteristics - they are specified at the very start of the model and do not change during the cab's path through it. These two attributes are used to separate the different groups of the cabs as the time spent in different processes depends on their values. Different cab models have dissimilar sizes and that affects the time in the painting processes, both base and coat, thus the attribute **Type** is used here to implement the time that correctly relates to the specific cab model. When it comes to attribute **Color**, which represents the color grouping as it was discussed at the end of Section 3.3.2, it has an impact on the probability that a cab is fully correct, or if it might need to be repainted, repolished etc, and therefore this attribute is used in the model to influence different probabilities for the cab's path.

3.4.3 Data usage in the model

The different values for attributes **Type** (6 different values representing the different cab model) and **Color** (3 different values representing the different grouping of colors) are assigned depending on the proportions in the data. A random number between 0 and 1 is generated and its value is compared to the proportions extracted from the data and the **Type** and **Color** are decided from that. For example, if 1164 out of 8488 observed cabs are from a Metallic color group, then there is a 1164/8488 chance that an *entity* will have **Color**-attribute value as "Metallic". The values for next 9 attributes (except the attributes concerning the time spend in the model) are assigned in a similar way, also depending on the proportion of the cabs that proceed to the specific processes retrieved from the data.

The use of the proportions from the data makes the model empirical and also as realistic as possible, with the only downside being a need of change of the applied proportions in the case of increase (or decrease) of the data sample. In that case all the numbers connected to the first 11 attributes and the distribution of their values would require to be recalculated and rewritten. This makes the current model quite data-dependant, but it is believed that a reasonable result can be retrieved from the model due to a quite large and therefore representative enough sample size.

When it comes to what time a cab spends in all different time-dependent processes (like painting, cooling zone, control etc) in the model, it is determined from a random value that is retrieved from a list with bootstrapped time data - as it was mentioned at the end of Section 3.3.2, the re-sampling method is used to represent the data in these processes. Using the bootstrapped data simplifies the choice of time that a cab spends in the process as it becomes unnecessary to investigate the distribution of the time in that process, but still provides a realistic implementation. The idea of the bootstrap method described in Section 2.1.2. When it comes to the stable processes like ovens and transportation between the time dependent processes, they are implemented as *Conveyor Systems* to have a constant time, as it was also mentioned in Section 3.3.2. The amount of physical spots in the processes, buffers and transport lines were carefully noted and thoughtfully implemented in the model in order to keep a realistic distribution of the cabs.

3.4.4 Elevators and transportation carts

Another important part of the model was implementation of elevators and transport carts. There are two main elevators that handle transportation of cabs between three different levels: 12-meter, consisting of two different buffer systems; 7-meter, the main part of the flow; and 0-meter, where the cabs arrive if they have any deviations. Where these elevators are placed in our model is shown in Figure 3.

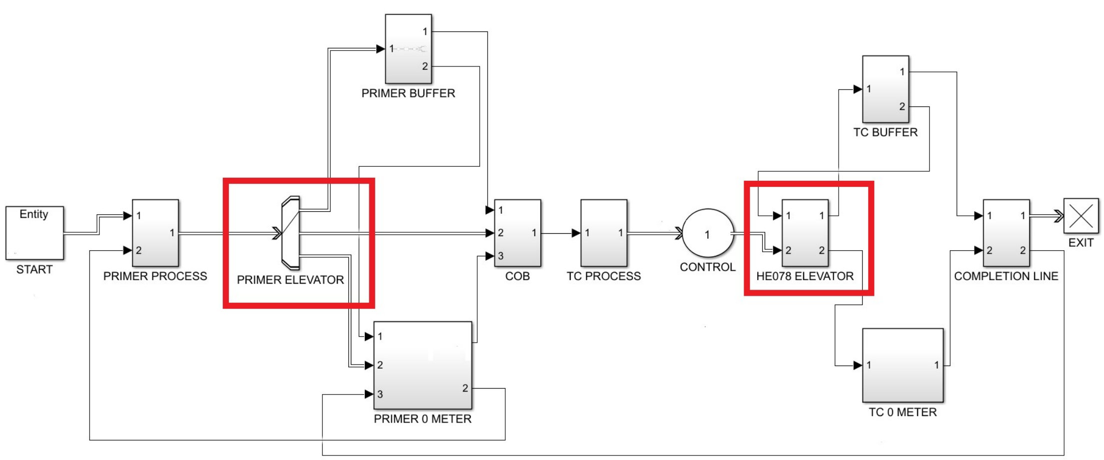


Figure 3: *Placement of the two main elevators in the model*

The first elevator is located after the primer coat processes, while the second elevator is located after the top coat process and control station. In the model they are implemented as *Entity Servers* to represent the time it takes to transport a cab. These two *Entity Servers* are connected to functions `switchHE138` and `HE078switch` respectively, which are described in the later Section 3.4.6. These elevators use attribute **Error** to make decision on a *entity's* path. The second elevator has a significant role as it also operates the transport of the cabs down from the top coat buffer, not only transporting them up. Therefore, a function `pickelevator` was implemented to prioritise which *entity* should enter the "Elevator" first - the one that is destined to go up or the one destined to go down. This function and its implementation are more thoroughly described in Section 3.4.6.

Other elevators which strictly work as a transport without any requirements to check a cab's destination, could not be implemented as *Conveyor System* (which worked well for the simple transportation), due to the fact that the elevators spend extra time to move to the original position. The same system was applied to transport carts which handle the transport of the cabs between two specific processes, where these transport carts move in-between. Thus, *Entity Servers* were used in the model where the *entities* spend a constant amount of time before moving on, to represent the process of an elevator/cart moving to receive a cab and then transporting it forward.

3.4.5 Buffers in the model

In the paint shop there are three significant buffers operating differently, which required a precise implementation in the model: primer coat buffer, color-optimization buffer and top coat buffer. Their roles in the paint shop were thoroughly described in Section 3.1 previously, and in this part their application and functionality in the model are discussed.

Primer coat buffer, which placement in the model can be observed in Figure 4, becomes a destination for the cabs in the case of color-optimization buffer (for the cabs without deviation) or the 0-meter level (for the cabs with deviation) being full. It consists of 52 buffer spots and is implemented in the model as a *Conveyor System* - the *entities* spend a constant time here and need to pass through all 52 places to exit the Buffer. Even though the time applied in a *Conveyor System* is constant, *entities* may not be able to leave directly if the next destination is full again, which recreates the variation in times spend here by the cabs.

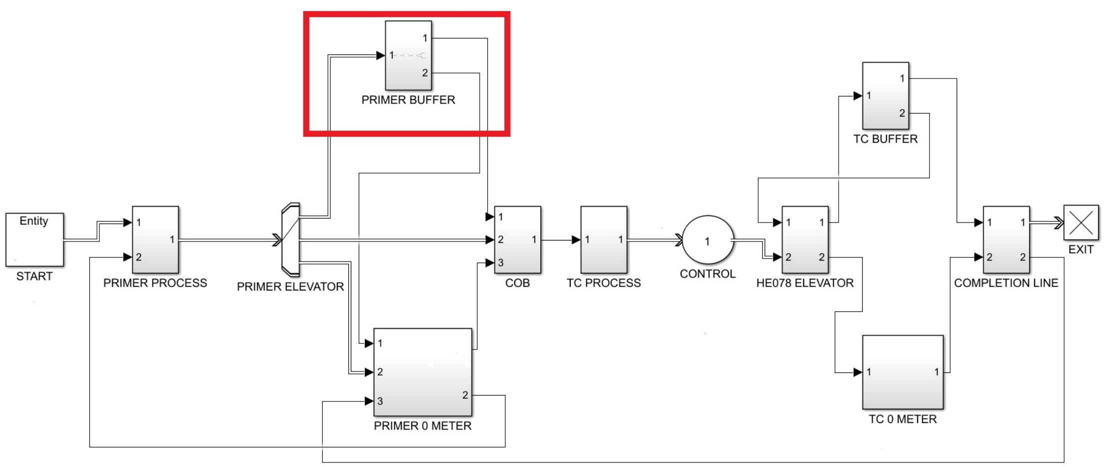


Figure 4: *Placement of the Primer Buffer in the model*

Another important buffer in the model is color-optimization buffer, which placement in the model can be observed in Figure 5. It is implemented by a *Queue* with 22 buffer spots from which the *entities* then separate to three different batches depending on the attribute **Color**. The batches are programmed to consist of 4 *entities* before they can proceed forward - in this way we replicate the "color sequence building"-mechanism which is the core of the COB.

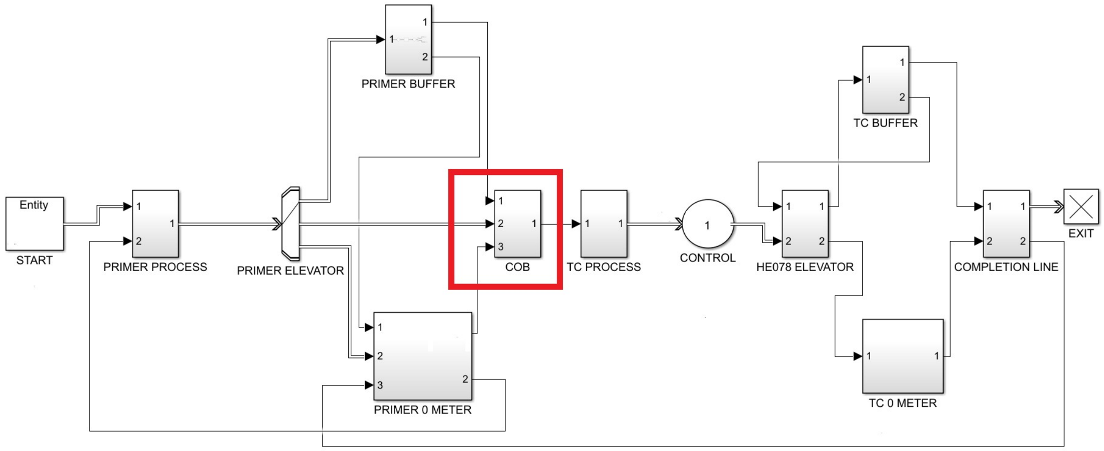


Figure 5: *Placement of the COB in the model*

When it comes to the top coat buffer, which placement in the model can be observed in Figure 6, its main characteristic is the fact that all cabs without a deviation have to proceed through it before they arrive at the final transportation line to then leave the paint shop. Another important detail is the fact that cabs with deviations and the ones without any are separated into two different lines which also later leads to them leaving the top coat buffer using different elevators. To have two separate buffer lines, in the model the top coat buffer first consists of one common *Conveyor System* which then separates into two different buffer lines (represented by the *Queues*) depending on the *entity* attribute **Error**. It resembled the reality quite accurately as the cabs firstly share the same path before separating and taking different passages before leaving for next destination.

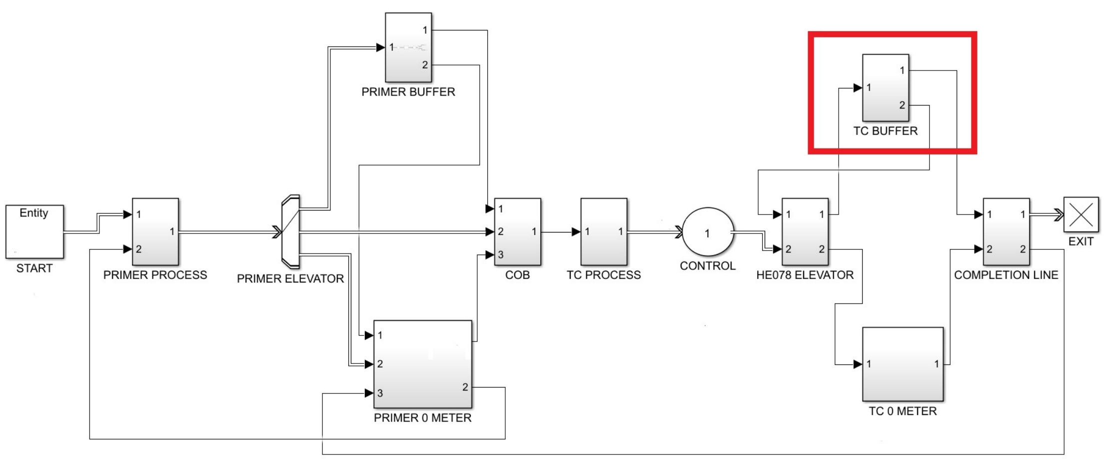


Figure 6: *Placement of the top coat buffer in the model*

In addition, to avoid over-flooding the top coat buffer, one *Gate* was implemented right before the top coat buffer to suppress the advancement of new *entities*. For that a function `gate_tl_buf` was used, and the implementation is thoroughly described in the next Section 3.4.6. Another detail of the top coat buffer is if the destinations ahead (transportation line forward to the exit in the case of no deviation or elevator to the 0-meter level in the case of deviation) are full, the cabs are sent back to the start of the buffer to go through it again. To handle that in the model a function `TLbuffert` is used, and its implementation is also described in Section 3.4.6.

3.4.6 Functions

3.4.6.1 switchHE138

There are two functions in the model that are connected to their respective main elevator, *Entity Servers*, to influence the path a cab takes in those elevators. The first elevator, whose location in the model can be observed in Figure 7, and the path after this *Entity Server* is implemented to depend on a function `switchHE138`. This function firstly checks if an *entity* has a deviation (Attribute **Error** is used here) and then secondly investigates if there is enough space in the next processes. Therefore, there are 3 inputs to the function: attribute **Error**, number of *entities* in the color-optimization Buffer and number of *entities* in the 0-meter level. Depending on the input values the functions return an output which correspond to which path *entity* will take. Therefore, it recreates the system where if intended next process is full, a cab proceeds to the 12-meter level to the primer coat buffer and then waits for the availability in those next processes.

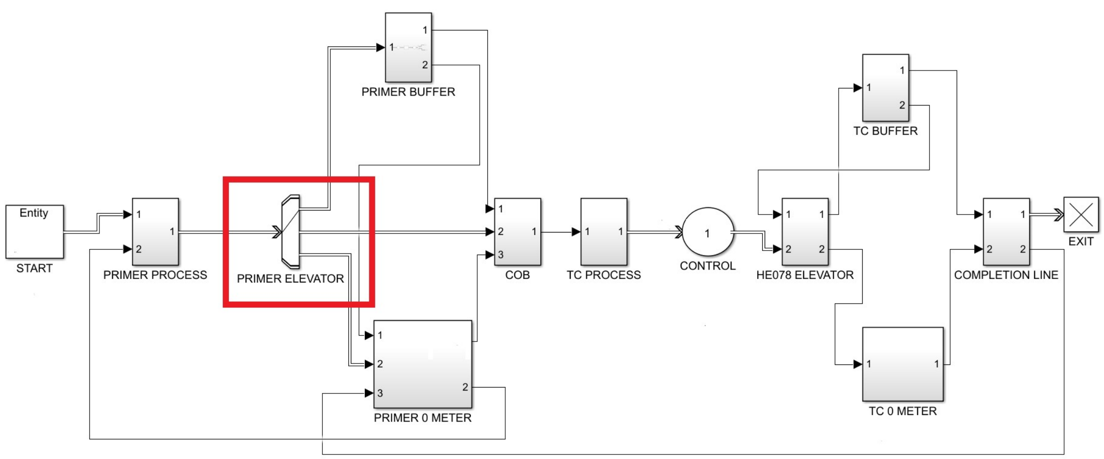


Figure 7: *Placement of the `switchHE138`-function in the model*

3.4.6.2 HE078switch

The second elevator, whose placement in the model can be observed in Figure 8, uses a function `HE078switch` to manage *entity's* path. Firstly, the function checks the attribute **Error** - if the "cab" does not have any deviations it is simply transported up to the top coat buffer. Otherwise, the function also takes number of *entities* present in the following processes: buffer line before the adjustment stations, buffer line before the improvement stations, a position on 0-meter level right before this elevator and buffer line at the 12-meter level. Thus, it investigates the workload in 4 processes that can be affected by the transport of additional *entities.* In a case when inputs gathered from 0-meter level processes indicate that these processes are full *entity* is sent up to 12-meter level. Therefore, this function manages *entities* without **Error** effortlessly as they are simply sent to the top coat buffer without any further checks, but in the case of **Error** the 0-meter level is investigated to avoid overfilling it.

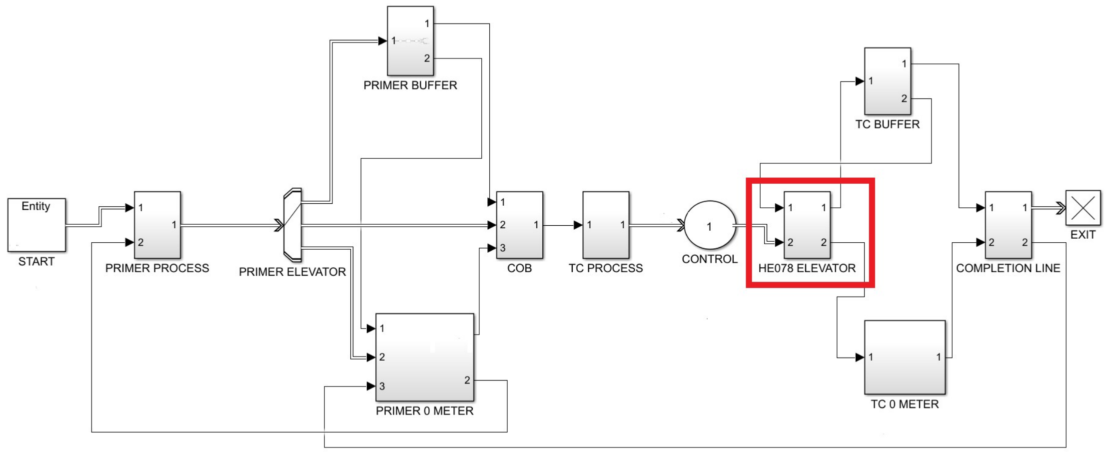


Figure 8: *Placement of the `HE078switch`-function in the model*

3.4.6.3 gate_tl_buf

Function `gate_tl_buf` is used to activate a *gate* in a case of top coat buffer getting overfilled, with the gate being located prior to the common part of top coat buffer. Thus, this function is called in the model at the start of the top coat buffer marked in Figure 9. This gate activates and blocks the flow as soon as the number of cabs in the buffer gets over 12. This implementation recreates the reality as accurately as it seemed possible - the fact that to release the pressure on the buffer the workers slow down the processes manually, thus making the decision depend on their personal judgement, makes this human interaction difficult to replicate in a more precise way.

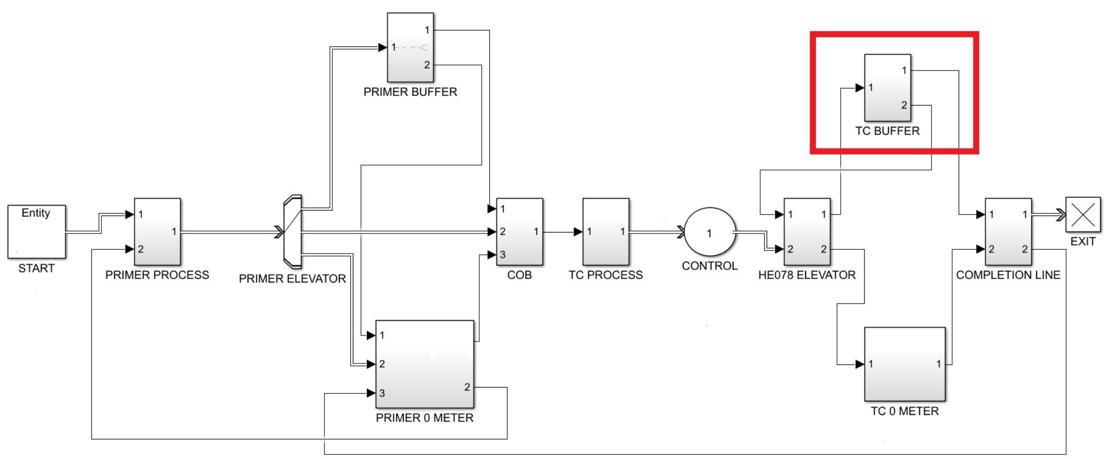


Figure 9: *Approximate placement of both `gate_tl_buf`- and `TLbuffert`-functions in the model*

3.4.6.4 TLbuffert

Function `TLbuffert` is activated when the *entities* depart from the common part of top coat buffer and try to enter the separated buffers. Thus, this function is called in the model at the end of the top coat buffer marked in Figure 9. The function firstly checks the attribute **Error** in order to identify the intended destination. In the case of no **Error**, the function checks if the buffer for the *entities* without errors, which is located ahead, is full - if it is then the *entity* is sent back to the start of the top coat buffer, if not the *entity* proceeds forward. Similarly, in the case of **Error**, the function check if the buffer for the *entities* with **Errors** is full - and the decision is made in the same way as for the other case. This recreates the real situations when the cabs just spin around in the buffer before they finally exit it, and also handles the overloading problem when it comes to the elevator which operates the transport to the lowest level.

3.4.6.5 NextSeed

Function `NextSeed` is implemented in order to make the processes in different simulations of the model completely random. The function generates and returns a Uniform Random Number, which then serves as a new seed for a Random Number Generation. This was required to solve the problem of MATLAB always starting with the same seed and therefore influencing random numbers to be the same between different simulations.

3.4.6.6 stampEntity

Function `stampEntity` simply returns the time in seconds since the start of the simulation. It is used together with *entity's* last 6 attributes to calculate the time an *entity* spends in specific different increments of the model. Firstly, the function is called at the start of the specific time interval and this time is assigned to that specific attribute. Then the function is called at the end of the chosen time interval and the attribute is updated to consist of the time in between these function calls. This provides an opportunity to analyse the times in different parts of the model and validate if they are reasonable.

3.4.7 0-meter level

0-meter level after the control station is a hugely important part of Volvo GTO's quality improvement system. If a cab has a deviation, it arrives here in order to be fixed. However, there are multiple stations here which serve different purpose and which initial destination a cab takes depends on the type of deviation, while cabs color also influences the path as some of the actions are not desirable for specific colors. Here the main processes are three adjusting stations, a grinding station located on the left side; and three improvement stations, one panel improvement station and a masking station which are located on the right side of this 0-meter level. All these processes are implemented as `Entity Servers` in the model. An overview of this level in Volvo GTO's internal system can be observed in Figure 10.

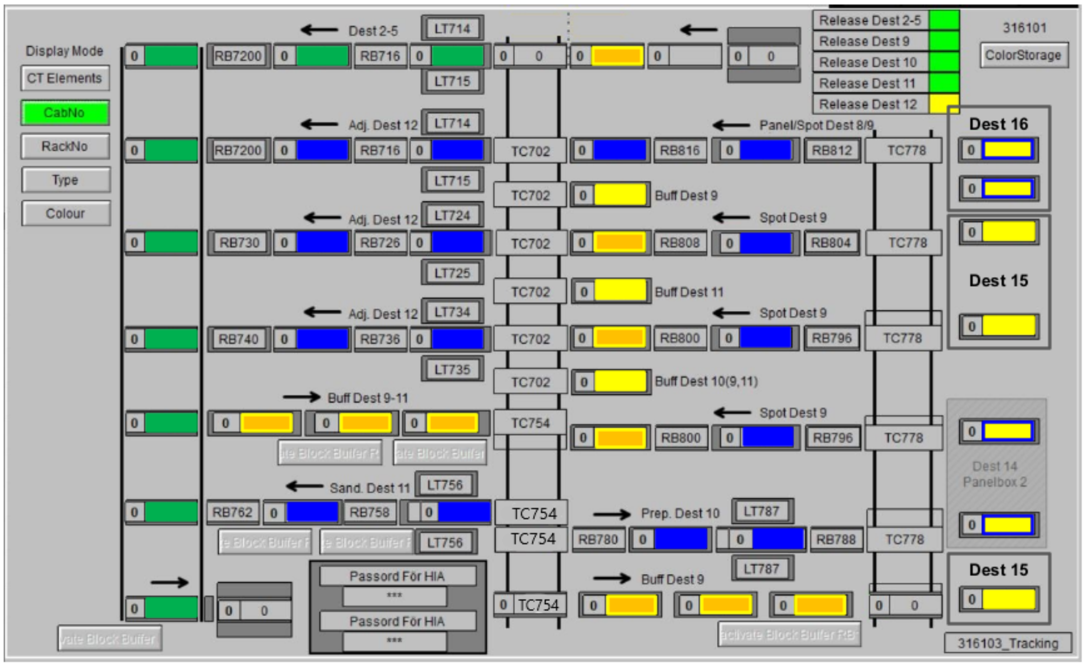


Figure 10: *A sketch of top coat 0-meter level where the cabs arrive in a case of deviation*

The blue color highlights the main processes, yellow - buffer places, yellow with blue edges - spots designed for repainting of cab's details but often used for buffer as well, green - line for cabs that are routed for the exit, orange - transport line which is often used as buffer as well. In Figure 10, the processes are marked with text which, for better understanding is described in further detail in Table 2 below.

Table 2: Main processes on the top coat 0-meter level

Name of the process in Volvo Trucks System	**Description of the process**
"Adj Dest 12"	Adjustment station
"Sand Dest 11"	Grinding station
"Panel/Spot Dest 8/9"	Panel improvement station
"Spot Dest 9"	Improvement station
"Prep Dest 10"	Masking station

In our model, the combination of different attribute values is used in order to determine where the cabs arrive to be adjusted and corrected. The layout of this 0-meter level in the model can be observed in Figure 11.

Firstly, the cabs arrive to a transport cart, called **TC702**, which has another transport cart next to it, called **TC754**. They have an important role of handling the transportation between the left and right sides of this level and in the model and they are implemented as `Entity Servers` with the explanation described in Section 3.4.4. **TC702**, handles transport to adjustment station or transport out when the cabs have no deviation, otherwise the cabs are sent to **TC754**. **TC754** handles the transport to the right side where the improvement and masking Stations are located, as well as transport to the grinding station which is located on the left side. Locations and therefore which transport these carts operate can be clearly observed in Figure 10. *Entity's* attribute **Error** is used to determine the initial destination. When *entity* comes out one of the main processes here, attribute values or the combination of attribute values, which are described in the Table 1 in Section 3.4.2, control the paths it will take next. As it was described in Section 3.4.3 these destinations depend on the proportions found in the data, and therefore *entities* can go through multiple different main processes here on the 0-meter level before leaving this level.

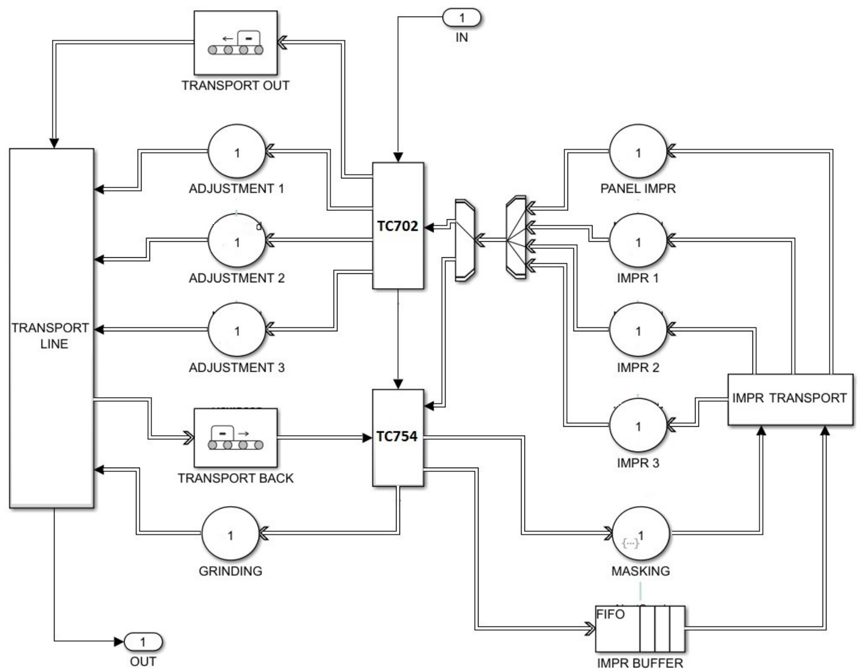


Figure 11: *A layout of top coat 0-meter level in the model*

If deviation is fully fixed here, which is marked by the *entity's* attribute **Error**, the *entity* proceeds to the final line for transport out of the model. However, if deviation is still existent, the cab may be sent back into the COB to go through the top coat painting processes again, or in other case to the truck-table for more adjustments.

3.5 Simulation Settings

3.5.1 Time

The unit of time used in the simulation is seconds. This is to more easily be able to vary parameters depending on changes in the factory as a multitude of production steps are as short as 30 seconds.

3.5.2 Schedule

The factory's general working schedule is 5 days per week, 24 hours per day. Thus, unless something unforeseen happens, what is currently in place over the weekend also remains there until the next workweek begins. This essentially means that the factory never starts from a clean slate but rather a spot where most positions are occupied. Due to this, the model requires a warm-up period of approximately a day before the real simulation can be started and data can be collected.

Because of the aim of simulating an ordinary week in the factory, the additional complexity of possible overtime during the weekend is not considered, as it can be seen as an anomaly.

3.5.3 Input Data

In order to run the simulation model there is a requirement of input data. The data can be divided into three parts, generation data, process times and probability functions. The generation data was gathered by ordering the starting points of cabs and calculating the difference between these. However, because there are stoppages in production from further down the line that are already implemented in the model, outliers caused by these have to be removed before the data can be used. The remaining two parts and the means for extracting them is explained in Section 3.3.2. Where process times are exported from R to MATLAB and probability functions are directly built into the corresponding entity servers in the model.

3.5.4 Running the model

The model was run for three instances with different purposes. Firstly, the model was run for 120 simulations with a simulation time of approximately 14 days and with a goal to determine the length of Warm-Up period. Secondly, two later simulation runs were performed in order to validate the model. One of them was a single simulation that run for 8 simulated weeks in order to provide a comparison with the empirical data, which also consisted of 8 weeks period. In the last simulation instance, the model was run for 204 simulations and the output was used to calculate the mean and median values from these simulations to further validate the model.

3.6 Warm-Up Period

As it was described in Section 2.3, a period in a simulation before the system reaches a steady state influences the output analysis and therefore a warm-up period is required to be selected. To formulate a reasonable warm-up period for our standard model, our Simulink model was run with simulation time of approximately 17 days. It was decided that our standard, or final, model would have simulation time of approximately 5 days which would represent one work week of production at Volvo GTO's paint shop. Starting from the first day of simulation, 5-day intervals were created with approximately 1 day increment in-between. Thus, the first interval represents simulation without any warm-up period, and then increment of 1 day slowly increases that period for each next interval, while the interval still keeps the same size of 5 days. Finally, to investigate the effect of warm-up period the median of total times was calculated for each 5-day interval and a plot of these values was created.

3.7 Model Validation

To validate the model the last 6 entity attributes - described in Table 1 - are used to calculate the time in model's different parts. where each attribute corresponds to each period.

Firstly, Total is the time for the whole process and is obtained from the attribute **Time**. For the remaining periods the time is calculated from where the previous period ended. Primer elevator is the first period which is the time from the start of the facility to the elevator after the primer paint process and is obtained from the attribute **Time_efter_GL**. The second period is labeled as TCemu which is the time until the end of the color optimisation buffer and is obtained from the attribute **Time_TL_EMU**. The third period is labeled Control which is the time until the control station past the top coat process and is obtained from the attribute **Time_Kontroll**. The penultimate period, labeled TT300 is the time until the cab reaches the transportation line towards the exit and is obtained from the attribute **Time_TT300**. Completion line is the final period which is the time until the cab exits the system and is obtained from the attribute **Time_OK**. For a clearer view of where each time period is registered see Figure 12.

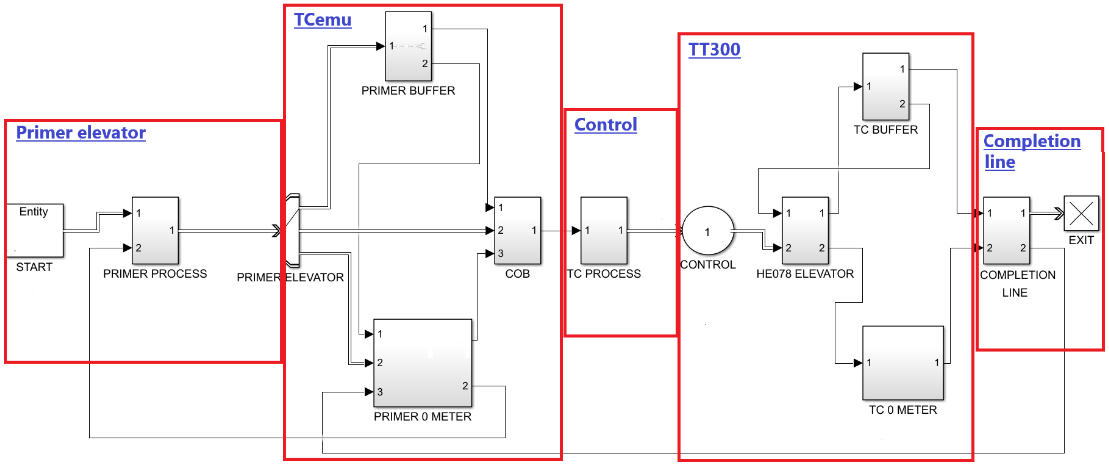


Figure 12: *Separation of the model to the 5 time intervals: Primer elevator, TCemu, Control, TT300 and Completion Line*

The mean and median values for times in these periods for any simulation could be compared to the mean and median values of empirical times, which are actual times it took for the cabs to proceed through the corresponding parts of the Volvo Paint Shop. This empirical data is retrieved from the dataset, which was described in Section 3.3.2, consisting of 8 different weeks of logged times for every cab in the system.

To analyse an approximate behavior of the model, one simulation was performed and the resulted times were then compared to the empirical times in a histogram together with their respective mean and median values. These plots were created for each period and furthermore all mean and median values for each period were presented in a table.

To validate the standard model further, it was run 204 times and the mean and median values for every simulation were calculated for each time period. These mean and median times were plotted in a histogram to observe their deviation. The mean and median values of these plotted times were added to the histograms together with the mean and median values from the empirical data.

3.8 Representations of Potential Process Modifications

This section goes over model and input data changes for three propositions of potential process enhancements. The results of the changes are then compared to the standard model over 204 independent simulations of each model. The comparison is done over a total of 6 different periods described in Section 3.7.

3.8.1 Changes to input data for the improvement station

At each improvement station, the time spent by the cab can be divided into two parts. The first part is where the physical repair is being performed. The second part is for heat being applied in order to incorporate the repairs so that new errors do not occur. The proposition is then to install new heating sources that cut the second part by 15 minutes. This is done in the model by altering the input data for the improvement stations by removing that amount of time from each value. However, as there are recorded times where a cab simply passes through the station due to various reasons, only times longer than 30 minutes receive the effect of the new heating sources.

3.8.2 Altering the probabilities at the control station

At the control station after the top coat painting process, a couple individuals scan the cab for defects. If a defect is detected, it is registered as a deviation and is generally sent for repair to the 0-meter level. However, due to no clear framework being in place for dealing with tiny defects, a percentage of cabs with deviations are sent straight to the completion line. Meanwhile, a multitude of registered times in the adjustment stations are lower than 100 which is not enough time for any repairs being performed there. These cabs are thus simply passing through the station with the only action applied being removal of the error classification so it can be sent to the completion line. The proposition is then to investigate the difference it would make to the system if a larger number of deviations were labeled as insignificant and those cabs thus being sent directly to the completion line. This is implemented by setting the probability of an error obtained in the top coat painting being overruled to 10% with the rest of the model remaining the same.

3.8.3 Adding an additional adjustment station

Due to the adjustment section being the most common destination for faulty cabs, these stations have a very high workload. A theoretical (but difficult in practice) proposition would therefore be to build an additional station there in order to spread the workload more and possibly improve the flow at the 0-meter level. This is implemented by altering the model to include an additional adjustment station identical to the three already in place. Additionally, the transportation line after the stations is also extended by one cab length to incorporate the output from the newly added station. This modification would result in the top coat 0-meter level of the model that can be observed in Figure 13.

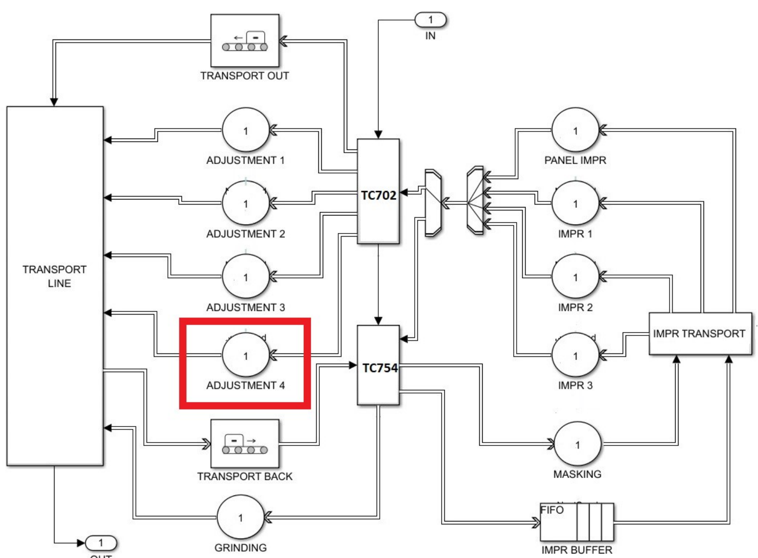


Figure 13: *A layout of top coat 0-meter level in the model with an additional adjustment station*

4 Results and Discussions

Our results consist of 4 different parts: firstly, the results of the Data Analysis are presented, which was an intensive part of the work before the actual model could be developed. Secondly, the result of calculation of Warm-Up period of the model is displayed, which then naturally leads to the results of the model validation, where the model performance is assessed. Finally, results are presented for three cases where the model was adjusted to examine three separate potential process enhancements. Succeeding the result of every section is a discussion of positives, reasonability and shortcomings along with connections to the real-life scenario.

4.1 Data Analysis

When analysing the empirical data as in Section 3.2 and 3.3.2, the purpose was to retrieve data for each part of the painting process in order to create an accurate representation of the paint shop in a simulation model. Because most parts of the system functions as a conveyor belt where the same speed is applied at all times unless a stoppage happens, the model is built with recursive elements. Essentially this means that the more individual processes are the cause of variation in the model while transportation is set to be constant but affected by the individual processes and thus taking on variation as well. Because of this, the empirical data is used in two different ways for presenting input data to the simulation model.

For transportation processes, the constant is set to be the general shortest time for cabs to travel through that process. The reason for it not being an ultimatum to use the time is due to inconsistencies in the registering of data and outliers thus having to be investigated before the actual shortest time can be decided. For the processes that are the cause of variation, data is resampled using bootstrap methodology from Section 2.1.2. However, with both natural variation as well as dependency on different factors being there, these had to be analysed first. Using ANOVA, the result obtained was that the specific model had a significant effect on both primer paint and top coat painting processes at a significance level 5 % which can be seen in Table 3 and Table 4. The model names represent three different sizes (1, 2 and 3) for each of the two models, FH and FM, which were mentioned earlier in Section 1.1.2. For the control process, using two sample Z-test showed for a significant difference between whether a deviation had occurred on the cab during the top coat process or not and can be seen in Table 5.

The color group of the cab only proved significant for time in the grinding section and more particularly for metallic cabs which can be seen in Table 6. However, it did prove to have an effect on both the odds of deviations occurring as well as where they were sent after passing through the quality control.

Approximately, 60 % of the cabs are in the normal color group, 25 % are in the uncommon group and 15 % are in the metallic group. In addition, both the uncommon and metallic groups had a higher probability of obtaining a deviation compared to the normal group. The general destination for these after passing the top coat control also differs between the groups and can be seen in Table 7. The main part that can be seen from this is that both the deviating groups have a higher probability of being sent to the more time-consuming sections compared to the normal color group. Also, the metallic cabs stand out in being sent far more often to panel improvement and grinding than both other groups. When looking at the destinations after a cab had visited any of the repair sections, the results varied. A summary of the most stand out information is presented here with a full list of destination probabilities available in the Appendix Table 16. Firstly, for all stations except grinding and masking, the majority of cabs having entered are sent directly to the completion line, with that option not being available for the mentioned two stations. Additionally, the adjustment section has the highest success rate of handling deviations for all colors and also has significantly different probabilities as a whole between them. In both improvement stations, the probabilities for uncommon and metallic cabs are not significantly different from each other and can thus be combined while the normal group still differs from the two. The differences end there though as no difference in destination probabilities could be derived for masking and grinding stations. Lastly, there was no difference in destination after a cab had been served by one station and then being sent to a different one, meaning that a cab can visit one station, be sent to another and then back to the first one again.

An additional note about destinations after the top coat control is that 2.1% of cabs with registered deviations are not sent to the 0-meter level but continue to the completion line directly without being sent down to the 0-meter level.

Table 3: The TukeyHSD result for comparisons of different cab models in the primer painting process. The notation **diff** is the mean difference between the two groups, **lwr** and **upr** are the lower and upper end points of the 95% confidence interval of mean differences between the group. The adjusted p-value for the comparison is notated as **p adj**

Model	**diff**	**lwr**	**upr**	**p adj**
FH3-FH1	27.09	21.11	33.06	0.00
FH3-FH1	42.80	36.31	49.28	0.00
FM3-FH1	-36.61	-44.05	-29.17	0.00
FM1-FH1	-15.03	-23.44	-6.62	0.00
FM2-FH1	-1.54	-10.15	7.07	0.99
FH3-FH3	15.71	11.25	20.18	0.00
FM3-FH3	-63.70	-69.46	-57.94	0.00
FM1-FH3	-42.12	-49.09	-35.15	0.00
FM2-FH3	-28.62	-35.84	-21.41	0.00
FM3-FH3	-79.41	-85.70	-73.12	0.00
FM1-FH3	-57.83	-65.25	-50.42	0.00
FM2-FH3	-44.34	-51.98	-36.69	0.00
FM1-FM3	21.58	13.32	29.84	0.00
FM2-FM3	35.07	26.60	43.54	0.00
FM2-FM1	13.49	4.16	22.83	0.00

Table 4: The TukeyHSD result for comparisons of different cab models in the top coat painting process. The notation **diff** is the mean difference between the two groups, **lwr** and **upr** are the lower and upper end points of the 95% confidence interval of mean differences between the group. The adjusted p-value for the comparison is notated as **p adj**

Model	diff	lwr	upr	p adj
FH3-FH1	21.80	16.10	27.50	0.00
FH3-FH1	31.04	24.86	37.22	0.00
FM3-FH1	-24.78	-31.87	-17.69	0.00
FM1-FH1	-11.31	-19.28	-3.33	0.00
FM2-FH1	-5.19	-13.38	3.00	0.46
FH3-FH3	9.24	5.00	13.47	0.00
FM3-FH3	-46.58	-52.06	-41.10	0.00
FM1-FH3	-33.11	-39.69	-26.53	0.00
FM2-FH3	-26.99	-33.83	-20.15	0.00
FM3-FH3	-55.82	-61.79	-49.84	0.00
FM1-FH3	-42.34	-49.34	-35.34	0.00
FM2-FH3	-36.23	-43.47	-28.98	0.00
FM1-FM3	13.47	5.66	21.29	0.00
FM2-FM3	19.59	11.56	27.62	0.00
FM2-FM1	6.12	-2.71	14.94	0.36

Table 5: Z-test results for time spent in the control process between cabs with errors and no errors

T–test result	Deviation mean	No deviation mean	Lower 95% CI	Upper 95% CI	P-value
Deviations vs No deviations	137.2	119.8	15.1	19.7	0.00

Table 6: The TukeyHSD result for comparisons of different color groups in the grinding section. The notation **diff** is the mean difference between the two groups, **lwr** and **upr** are the lower and upper end points of the 95% confidence interval of mean differences between the group. The adjusted p-value for the comparison is notated as **p adj**

Color Group	diff	lwr	upr	p adj
Normal-Metallic	-1222.27	-2309.20	-135.33	0.02
Uncommon-Metallic	-1930.00	-3003.85	-856.15	0.00
Uncommon-Normal	-707.73	-1647.57	232.10	0.18

Table 7: The probabilities of which station at top coat 0 meter the cab will be directed to for each color group, given that cab has obtained a deviation

Color Group	Adjustment	Improvement	Panel Improvement	Grinding	Masking
Normal	75.7%	18.4%	4.6%	1.2%	0.10%
Uncommon	67.8%	26.2%	3.8%	2.0%	0.20%
Metallic	56.6%	28.9%	7.7%	6.5%	0.30%

4.1.1 Discussion of Data Analysis

Given that the time spent in transportation processes depends on how cabs are situated before and after the process along with the decision making of process controllers, the model has been built in a specific manner. In order to both represent the short, normal state times and longer times based on interaction with other processes for transportation, times for these are set to be the shortest recorded. This means that in order to obtain some of the longer times, influence has to come from further ahead in the process. In short, this is an attempt to recreate the role of process controllers but without the general feedback for other parts in the system. Thus, the result of this methodology is that evasive actions in order to avoid exceedingly long stoppages are sometimes not taken. This means that while the system continues to function properly at all times, for situations where the system is overfull, times may be exceedingly long for a period of time.

For resampled times, overly excessive outliers have been removed but, there is still a dependency on how the data has been collected and whether the cab was actually meant to be served in that process or not. This issue also has an effect on the pathing probabilities at 0-meter level as the probabilities are solely based on empirical information from the data.

Finally, the comparison for the control process was only made between deviation against no deviation with no dependency on the number of deviations. The reasoning for this was that additional deviations can occur later on in the process so when analysis was made for time based on total number of deviations, no clear result was obtained. Additionally, as the control is operated by humans, generally an extra thorough check of the cab is performed if any defect is detected. This means that even if no additional defect is detected, approximately the same amount of time is spent observing the cab as if another defect would have been found during that extra check.

Generally, most dependencies from within the factory have been identified and quantified for representation in the model. Shortcomings here would mainly be colors being combined into one of three color groups and identification of when human labour is available or not. Given this, it seems reasonable that the model should be able to recreate the painting process approximately.

4.2 Determining the Length of the Warm-Up period

As was described in Section 3.6, intervals of 5-day simulation time were created with approximately 1 day increment in-between to investigate the effect of the warm-up period. Calculated median values of total times for each such interval can be observed in Figure 14. Furthermore, the mean and median values were calculated for the very first 5-day interval, which represents the model without any warm-up period at all. These mean and median values compared to the empirical values are presented in Table 8.

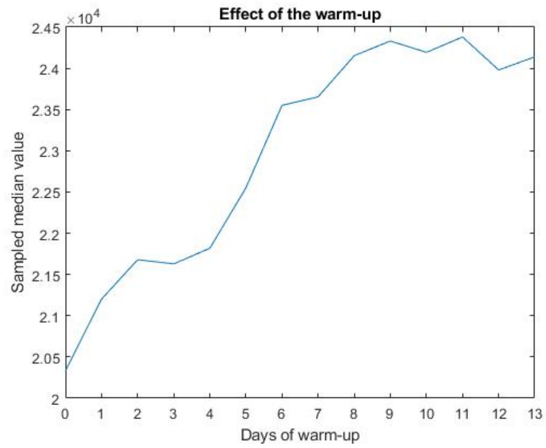


Figure 14: *The effect of warm-up with approximately 1-day increment*

Table 8: Mean and median values for total time in the model without warm-up and empirical data

	Mean	**Median**
Data-based	35005	30828
No warm-up model	24273	20331

From the figure it could be noticed that median value quickly increases as the simulation time increases which could be expected as the model fills with more and more objects. Furthermore, mean and median values of the no-warm-up model are considerably lower than the empirical values, which can be explained by the fact that the model is empty when a simulation starts and therefore there is no steady state during the beginning period. It can be observed that 8 days since the start of the simulation would be a reasonable warm-up period as the increase of median value stagnates and stays on a similar level afterwards.

The standard model was run and resulted in the output that consists of 5 simulated days, which start after this 8-day warm-up. The mean and median values for the total time it takes to go through the whole system of this standard model were calculated. They were compared with the mean and median values for the actual total time in the system calculated from the data and these values are presented in the Table 9.

Table 9: Mean and median values for total time in the model with warm-up and empirical data

	Mean	**Median**
Data-based	35005	30828
Final model	36607	24338

From the table we can observe that the mean and median values differ, the standard model has lower median values, however it's mean value is a bit larger than its empirical counterpart - it might indicate a larger variance in the model.

4.2.1 Warm-Up discussion

The idea with warm-up period was the fact that in reality, if you would start measuring one specific cab's time in the Paint Shop from the start, there are already many cabs in all the processes in the whole system from the previous days and weeks of production. This obviously affects both the path and time of the cab in question, and therefore the output of our model at the start of the simulation becomes unreliable as the model is completely empty in contrast to the Volvo's factory system.

As the objective of the model was defined to simulate one working week at the Volvo Truck Paint Shop, the final output was aimed to consist of approximately 5 days. Thus, for the calculation of warm-up period it seemed very reasonable to just compare multiple models, with the same total 5-day simulation time, with the difference being the length of the warm-up period. When it comes to the choice of the measure, the median was deemed suitable as it is more stable than a measure of mean considering the variation and randomness in the analysed process. Therefore, when we try to find the moment when this median value stagnates - we stabilize the measure which is already quite stable.

In the shown figure we could see that the median values are very small at the start, but increase as length of the warm-up period increases. At 8-days of warm-up we can see that the median values stagnate and stay approximately on the same level - which therefore led us to choose that length for the warm-up period in our final, standard model. Furthermore, standard model's mean and median values seem to be quite similar to the empirical values, which indicates the reasonable choice of warm-up period.

4.3 Model Validation

Here the results of validation of the standard model are presented, where firstly, times from one random simulation are compared to the empirical times, and secondly, mean and median values for multiple simulation results are plotted and analysed.

4.3.1 One simulation compared to the empirical data

Total time from one random simulation is compared to the empirical total time, and this plot can be observed in Figure 15, Mean and median values for each case are also plotted in the histogram, and these values are presented in Table 10.

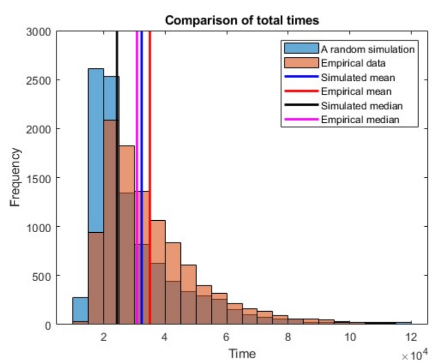
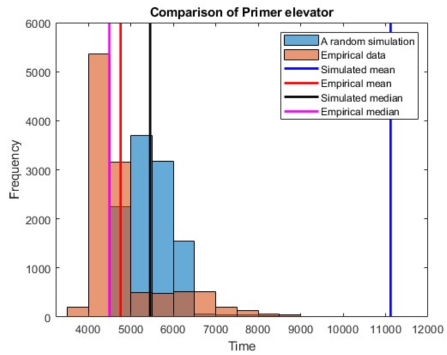


Figure 15: *Total times for a random simulation and empirical data*

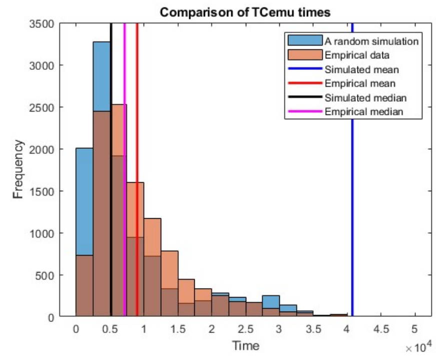
Table 10: Mean and median values of Total times for a random simulation and empirical data

	Mean	**Median**
Simulated	33987	24678
Empirical	35005	30828

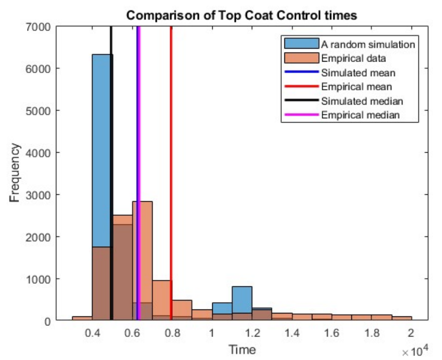
Given the results above it can be noted that the total times did not differ especially much which offers a positive indication. However, for a deeper understanding and insight in the variation of times in the model, a closer look at the different parts of the total time would be desirable. For that further analysis, the similar histogram plots of times in 5 different time intervals for one random simulation and the empirical data were created. Figure 16 shows the histograms of times in 5 different time intervals. Furthermore, the figure also contains the mean and median values for each case which are then presented in Table 11.



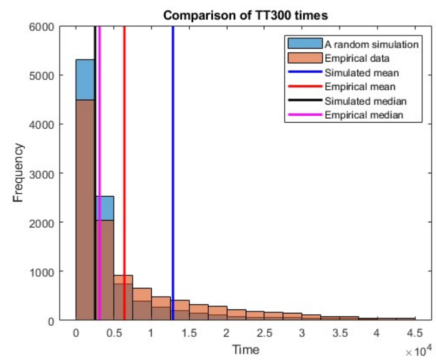
(a) Primer Elevator times for a random simulation and empirical data



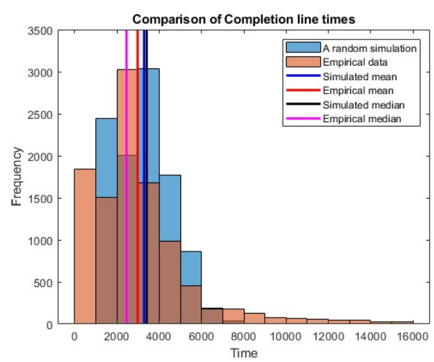
(b) TCemu times for a random simulation and empirical data



(c) Control times for a random simulation and empirical data



(d) TT300 times for a random simulation and empirical data



(e) Completion line times for a random simulation and empirical data

Figure 16: Comparison of times for a random simulation and empirical data for the different periods: Primer Elevator, TCemu, Control, TT300 and Completion line

Table 11: Mean and median values of times in 5 different period for both a random simulation and empirical data

Time interval	**Value**	**Simulated**	**Empirical**
Primer Elevator	Mean	11119	4755
Primer Elevator	Median	5452	4490
TCemu	Mean	46942	9012
TCemu	Median	5328	7159
Control	Mean	6362	7939
Control	Median	4936	6338
TT300	Mean	13879	6410
TT300	Median	2522	3120
Completion line	Mean	3183	2981
Completion line	Median	3215	2458

A very high variance can be noticed in some periods indicated by noticeable mean values, while median values in these processes were still close to their empirical counterparts. This suggests that periods in Primer elevator, TCemu and TT300 are quite significant in the model and might cause cases with extreme values.

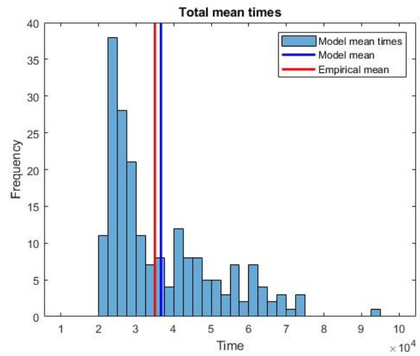
4.3.2 Mean and median of times from 204 simulations

Table 12 contains information from 204 simulations of a standard week in the facility for the standard model and also information from the empirical data. The information is split into different segments of the painting process, namely the total time and its 5 intervals: Primer elevator, TCemu, Control, TT300 and Completion line, with the values being of the mean and median values.

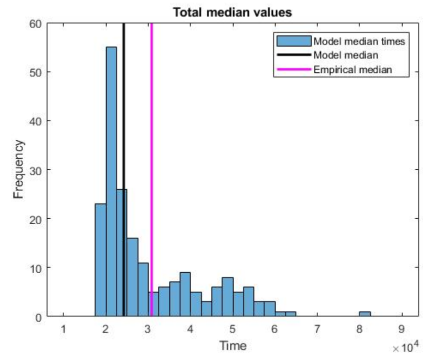
Table 12: Mean and median values from the standard model and empirical data for each period

Source	**Standard Model**	**Standard Model**	**Empirical Data**	**Empirical Data**
Values	Mean	Median	Mean	Median
Total	36607	24338	35005	30828
Primer Elevator	8539	5482	4755	4490
TCemu	24588	5725	9012	7159
Control	6784	4935	7939	6338
TT300	10777	2260	6410	3120
Completion Line	2944	2643	2981	2458

Graphical representation of both mean and median values for the total time is presented in Figure 17, together with the empirical mean and median presented in Table 12 above.



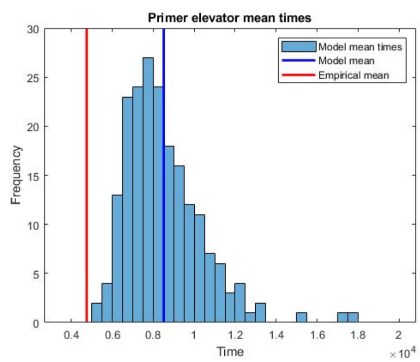
(a) Histogram of mean values for total times



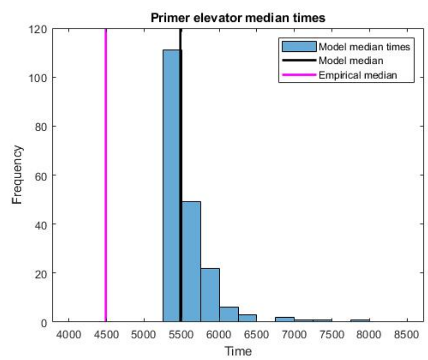
(b) Histogram of median values for total times

Figure 17: Histogram of mean and median values for total times

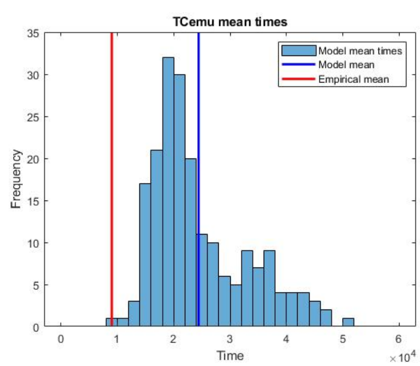
Further the mean and median times for each partial time interval are plotted in a histogram, together with the corresponding empirical values presented in the Table above. These histograms are shown in Figures 18 and 19.



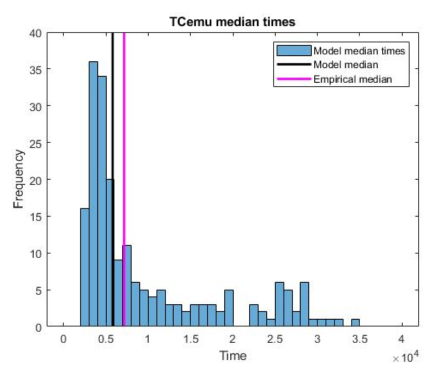
(a) Histogram of mean values for Primer Elevator times



(b) Histogram of median values for Primer Elevator times

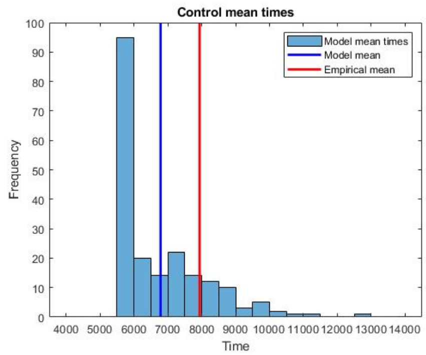


(c) Histogram of mean values for TCemu times

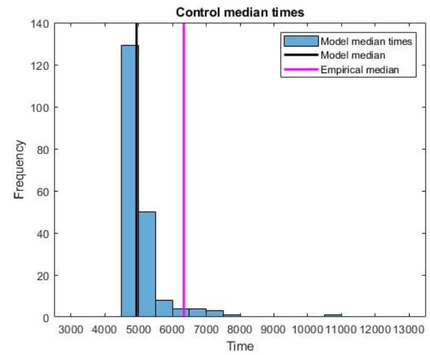


(d) Histogram of median values for TCemu times

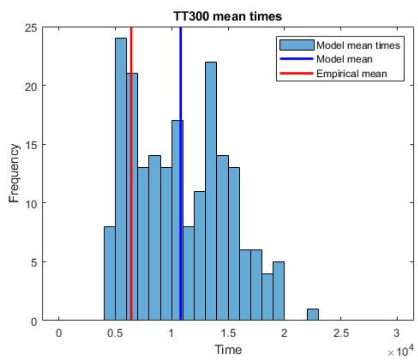
Figure 18: Histogram of mean and median values for Primer elevator and TCemu times



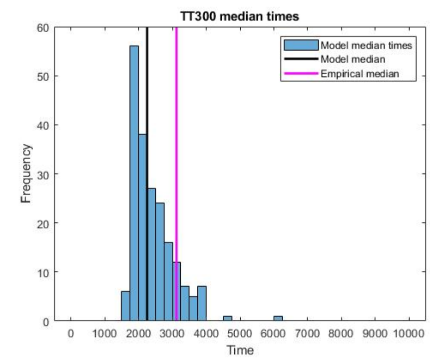
(a) Histogram of mean values for Control times



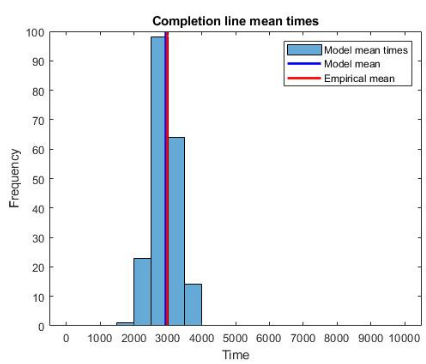
(b) Histogram of median values for Control times



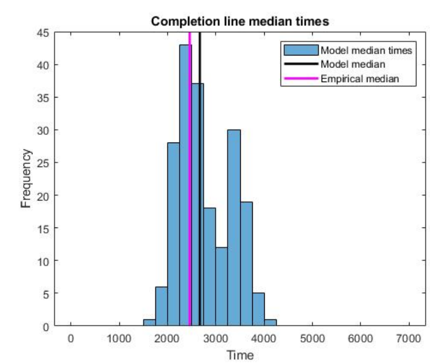
(c) Histogram of mean values for TT300 times



(d) Histogram of median values for TT300 times



(e) Histogram of mean values for Completion line times



(f) Histogram of median values for Completion line times

Figure 19: Histogram of mean and median values for Control, TT300 and Completion line times

Given the results above it is clear that there is quite large variation in some processes even though the total time mean and median appear to not deviate as much. As it was noticed in the previous validation section, mean values in periods Primer elevator and TCemu differ quite a bit from the empirical values. Completion line can be identified as the most stable process which could be expected as all the objects here proceed through the same positions in a quite small period of time.

4.3.3 Discussion of Model Validation

While there are clearly differences between the empirical data and the simulation model, the fact that the distributions of values obtained in Figure 15 and Figure 16 seem to follow the empirical distributions with slight discrepancies does bode well for the model. Total mean values from Table 12 show that over time, cabs generally occupy spaces in the system for the same amount of time. This means that while the median time in the system is lower for the model and thus more cabs go through the system faster than expected, the remaining cabs are still in the system. As the model is built with the purpose of queues forming to slow down transportation as mentioned in Section 4.1.1, this should mean that the aim of recreating both long and short transportation times is successful. Then, realistically, if an attempted alteration does not have an effect on the model, it should not have a big effect on reality either as both highs and lows are represented along with the average time spent in the process being the same.

Furthermore, Completion line estimates are eerily similar which shows that as cabs leave the system at the same rate, some stability will always be in place. This means that while there is an excess in shorter times, these cabs will still have to go through the Completion line for as long as anyone else, stopping cabs from continuously racing through the system and effects on the system should still apply to these. Additionally, when looking at the distributions for a random simulation in Figure 16 it is important to note that randomness has been introduced to the system which means that elements that have been represented in the simulation but to a lesser extent than in the data, may become more present in another simulation, or vice versa. The issue of outliers also has its place to discuss here as the simulated mean for Primer Elevator and TCemu times does not make sense in the context of Figure 16. Due to the functionality of Simulink, there can be occasions where the software decides to exclusively pick from one side of a crossroad which then may lead to cabs becoming stuck in the system for a long time. This is most prevalent after an extended stoppage where a multitude of cabs are lined up to continue through the process so there will be a continuous flow of cabs passing through the system and thus, some cabs can get stuck behind this sequence. A decision regarding empirical data was also taken to remove exceedingly large outliers as these can occur due to various reasons including time being recorded over the idle weekend. This was made in order to have more of an ordinary production week to compare the results to but, does inhibit the comparison of outliers. Given this, median values should be given a larger role in the comparison as these are more stable, with mean values as a good compliment as it becomes a measure of how long cabs take up a space in the system.

Bearing in mind all mentioned above, Figures 17-19 show that while discrepancies are present, that was always to be expected due to the complexity of the system as multiple process controllers are employed with the purpose of simply ensuring that the process functions somewhat properly. In the absence of process controllers, extreme values would be ever more common and would most likely occur throughout the process. Thus, the fact that simulated median and mean values for all periods of the process seem to stick close to their counterpart in the empirical data while switching sides throughout so that they sometimes cater towards being larger and sometimes smaller can be seen as a positive outcome.

Conclusively, the model shows positive signs of inheriting the elements of the real-life process while introducing random elements that ensure that changes implemented to the system have multiple opportunities to make its effect noticeable. This allows for the possibility to implement changes and investigate how different parts of the process might be affected by these to predict which possible countermeasures can be taken in the real process in order to reap the benefits of an alteration. Given the randomness and discrepancies in the model however, it is important to look at more than just average values and instead look at how these are distributed and how outliers may or may not have a real connection to the process.

4.4 Potential Process Modifications

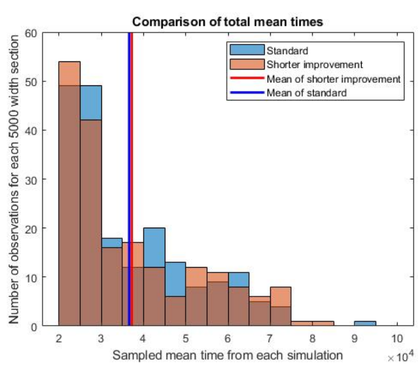
Here the results of the different modifications are presented and compared to the results of the standard model in order to visualise the effect these have on the system. Comparisons are based on sample values for each simulation so when mentioning a single mean, it corresponds to the mean value of all mean values obtained from the simulations. Likewise, when mentioning a single median, it corresponds to the median value of all observed medians from each of the simulations. SD of Means is the standard deviation of all observed means while SD of Medians is the same but calculated on all observed medians.

4.4.1 Modification 1: Shorter Improvement Times

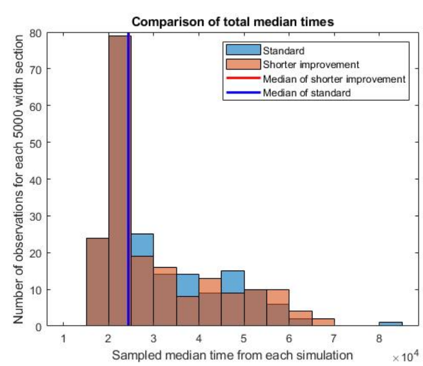
Table 13 contains information from 204 simulations of a standard week in the facility for the standard model and the model based on modification 1. The information is split into different segments of the painting process, namely Primer elevator, TCemu, Control, Completion line and the total time, with the values being of the mean and median values along with the standard deviation of these. Graphical explanations for the most important and varying periods, total, Primer elevator, TCemu and TT300 can be obtained from the histograms in Figure 20 and Figure 21, with the information in Table 13 being deemed enough to evaluate the two remaining periods.

Table 13: Comparison of simulated means, medians and standard deviations between the standard model and the model based on modification 1 with shorter improvement times

Model	**Standard**	**Modification1**	**Standard**	**Modification1**	**Standard**	**Modification1**	**Standard**	**Modification1**
Values	Mean	Mean	SD of Means	SD of Means	Median	Median	SD of Medians	SD of Medians
Total	36607	37229	14387	15724	24338	24443	12303	13052
Primer Elevator	8539	8890	1925	2261	5482	5497	986	1110
TCemu	24588	25541	8633	8833	5725	6081	8454	8521
Control	6784	6882	1282	1439	4935	4939	643	712
TT300	10777	10820	4236	4662	2260	2502	658	730
Completion Line	2944	2971	378	380	2643	2648	545	528

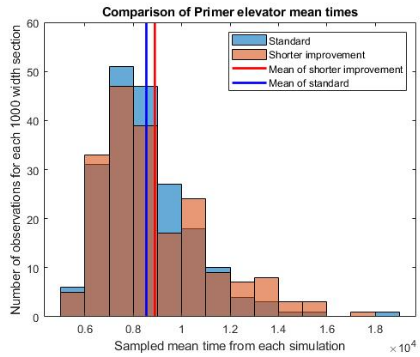


(a) Total time in the system mean

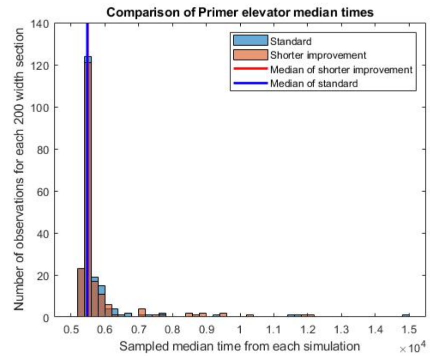


(b) Total time in the system median

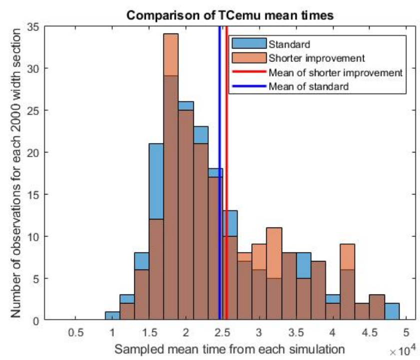
Figure 20: Comparison of mean and median times for the total process time between the standard model and the model for modification 1



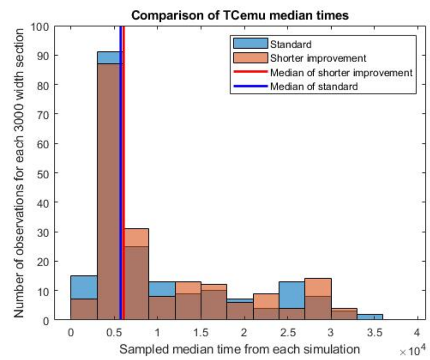
(a) Primer elevator mean



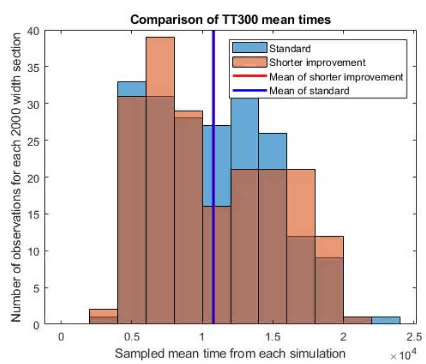
(b) Primer elevator median



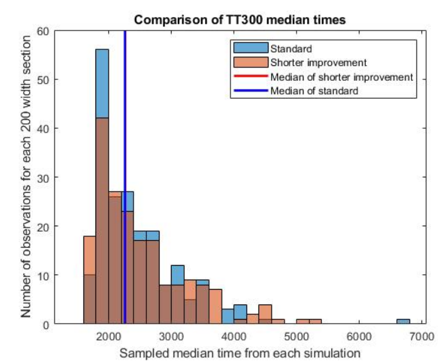
(c) TCemu mean



(d) TCemu median



(e) TT300 mean



(f) TT300 median

Figure 21: Comparison of mean and median times for the time periods of Primer elevator, TCemu and TT300 between the standard model and the model for modification 1

Given the results above it is clear that the implemented change does not have a big impact on the system at any stage of the process. The change does seem to have an effect on how the values are spread out over time though, in particular for TT300 as the values seem to have been spread out more from the center, something that is also implicated by the measures of standard deviations there.

4.4.2 Discussion of Modification 1

From Table 13 it can be seen that almost all observed values have increased from the implemented change of a more efficient heating lamp. While the median values have been affected the least, the effect made on all estimates of standard deviation is interesting. Figure 20 and Figure 21 give a better explanation for this occurrence. Generally, the mean and median values have remained the same which would tell us that the effect of the change is negligible. However, these figures show that the mean and median values have been spread out very differently. In comparison to the result from the standard model, the change has seen an increase in values both above and below the mean for almost all periods. This is represented by more blue being present surrounding the mean and median lines while more red is present as values move further away from them.

A short explanation for the result is that the model has become less steady with shorter time spent at improvement meaning that the transportation wagons at 0-meter level have less idle time and thus, there is a higher risk for unwanted stoppages when a multitude of cabs obtain deviations in short succession. A more in-depth explanation and how this change would affect the real process has a couple more factors to take into account however.

The first factor is the one mentioned in Section 4.1.1 about how the model methodology may have an effect on the scenarios where the model is overfull. As the change has not exclusively seen an influx of longer times but also in times shorter than the middle, there might be more of an effect of the change that was first believed. The values investigated are also only mean and median values for each simulation, meaning that the effect on individual times is not shown. Mean values are also highly susceptible to extreme values which means that for there to have been an influx in shorter mean times, the change does show promise for making the process more efficient. When looking at the practical implementation of this modification, there are some uncertainties. With improvement stations having both active and idle time in the process, there is the possibility for employees stationed there to serve one cab while the other one is being treated by the heating source. A decrease in idle times could then result in the current personnel not being sufficient to fill the needs at the same tempo as what is currently in place.

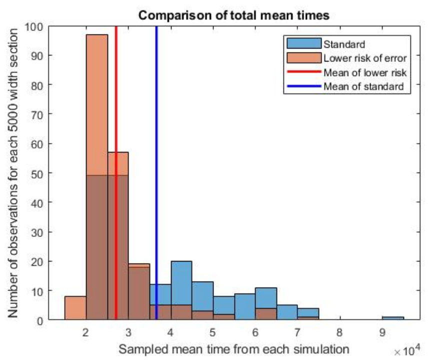
Taking all of the above into account, the actual effect of this change is difficult to estimate with the model. It does however suggest that given no other modifications it does not seem reasonable to spend money on changing the heating lamps with no other change in place for the process. This, as the general time for a cab to go through the system has seen negligible changes and no positive change in workload for personnel at 0 meter would be caused by this. If the change could be combined with a new framework for how personnel act at the improvement stations, there could be use in that sense but as that would require rescheduling of tasks, no information regarding this can be obtained from the model.

4.4.3 Modification 2: Lower Risk of Error

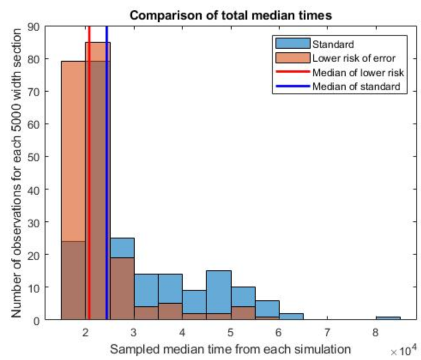
Table 14 contains information from 204 simulations of a standard week in the facility for the standard model and the model based on modification 2. The information is split into different segments of the painting process, namely Primer Elevator, TCemu, Control, Completion line and the total time, with the values being of the mean and median values along with the standard deviation of these. Graphical explanations for the most important and varying periods, total, Primer elevator, TCemu and TT300 can be obtained from the histograms in Figure 22 and Figure 23, with the information in Table 14 being deemed enough to evaluate the two remaining periods.

Table 14: Comparison of simulated means, medians and standard deviations between the standard model and the model based on modification 2 with lower risk of error

Model	**Standard**	**Modification2**	**Standard**	**Modification2**	**Standard**	**Modification2**	**Standard**	**Modification2**
Values	Mean	Mean	SD of Mean	SD of Mean	Median	Median	SD of Median	SD of Median
Total	36607	27073	14387	8483	24338	20736	12303	6875
Primer Elevator	8539	8117	1925	1364	5482	5461	986	350
TCemu	24588	19199	8633	5311	5725	4243	8454	4521
Control	6784	6097	1282	758	4935	4904	643	334
TT300	10777	7196	4236	3248	2260	1746	658	311
Completion Line	2944	2647	378	277	2643	2293	545	376

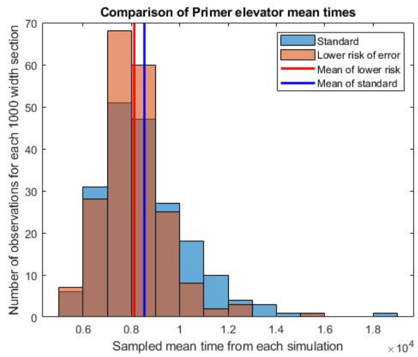


(a) Total time in the system mean

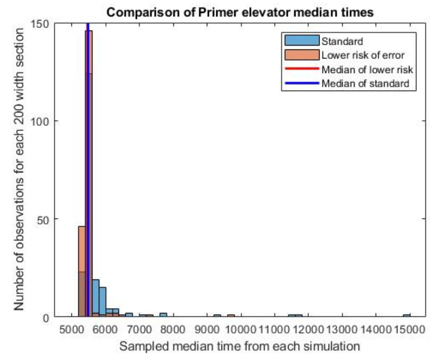


(b) Total time in the system median

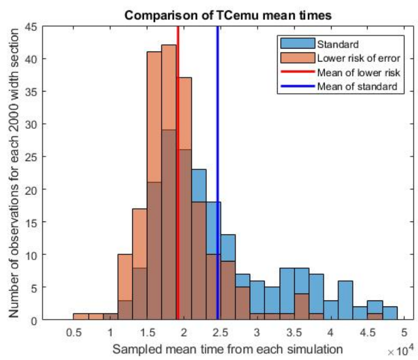
Figure 22: Comparison of mean and median times for the total process time between the standard model and the model for modification 2



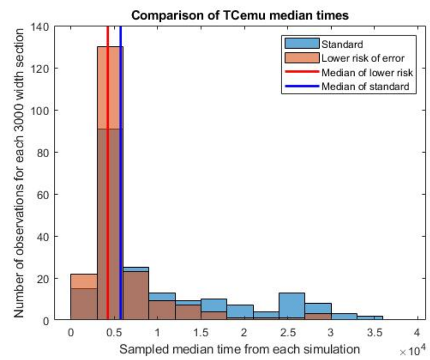
(a) Primer elevator mean



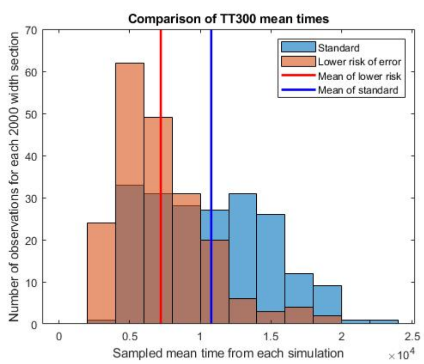
(b) Primer elevator median



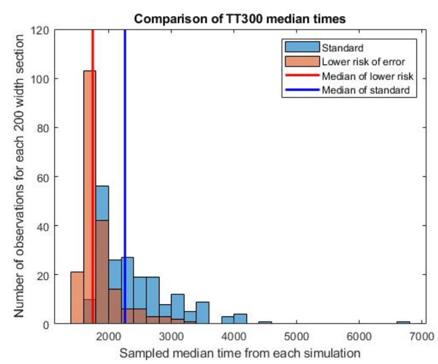
(c) TCemu mean



(d) TCemu median



(e) TT300 mean



(f) TT300 median

Figure 23: Comparison of mean and median times for the time periods of Primer elevator, TCemu and TT300 between the standard model and the model for modification 2

Given the results above it is clear that the implemented change has a drastic effect on the system as essentially all mean and median values for each part of the process have decreased by a lot. The system has also become more stable as all measures of standard deviation have decreased.

4.4.4 Discussion of Modification 2

From Table 14 it can be seen that almost all observed values have decreased drastically from the implemented change of sending a higher percentage of cabs with registered deviations directly to the completion line. While all values have been lowered, the most interesting to note is that standard deviation values have seen a drastic decrease everywhere which shows that the system has become more stable. The values of standard deviation have also decreased for the two generally least affected processes, Primer Elevator and Completion Line. Interestingly, these are the first and final parts of the system and are periods that in reality, generally are affected by different sections of the facility at Volvo than what is taken into account in this model.

Figure 23 shows that while the median and mean lines have all been moved back, the general change in times has not been around the lines of the standard model. Instead, the difference is mainly found in values above the lines being far less common while those below have increased in frequency by a lot. Essentially this means that the histogram has been shortened and thus more stability has been introduced in the model by the proposed change.

As mentioned in Section 4.4.2, the methodology used to replicate transportation times may have an effect on how the implementation actually affects the system. The stability obtained from a lower number of cabs being sent for repair seems to have almost completely removed the issue of the system becoming overfull. While this may be something that generally lowers the dependency on process operators, there is also the risk of the model losing its connection to reality. As transportation times are set to both replicate highs and lows, if there are barely any stoppages in the process, the risk is that transportation hardly ever gets slowed down. However, as the way transportation is slowed down in reality is based on stoppages and possible upcoming stoppages in the process, if the effect is similar there to the one in the model, there is the possibility of it becoming the new normal state.

There are still some issues to be dealt with regarding the implementation of this proposal and its effects on reality.

First and most importantly, simply sending cabs with defects to the completion line will have drastic effects if those defects actually have to be repaired and as we do not have information about how big the issue actually is, this part is not looked into in the model. From discussions with multiple personnel at Volvo however, the general consensus seems to be that a lot more cabs are being sent down with defects than what people higher up in the hierarchy believe. It also has been established to be a common occurrence that cabs with very slight defects are sent for repair but are simply sent through to completion directly instead as the risk of causing further defects is very high and the observed one is simply something that will occur during transportation to the customer anyway. The number 10 % is also arbitrary and has no scientific reasoning behind it besides being larger than the current 2.1% while still not being large enough to the point where a lot of real errors are being passed upon.

Secondly, while the increased percentage is based upon multiple far too short values being present in the time vector for the adjustment stations, no modification has been made to this vector to account for the cabs being sent to the completion line. While this seems like a big issue, there are a few reasons for it perhaps not being as impactful as first thought. Firstly, it is difficult to estimate how many cabs there are that actually get served at 0 meter as the processes are handled by actual humans which means that if no one is available, times may still be high despite no action being taken regarding the cab. Secondly, it is not exclusively when deviations are ignored that short times are recorded, but also when the system is full and a cab destined for a different station is sent through a process to act as an additional buffer spot. With no knowledge of which times belong to which action, a decision on which and how many times should be removed becomes almost arbitrary. Lastly, given that less cabs are sent down for repairs, the personnel there should be able to act more quickly as they should be able to take breaks at more appropriate times. This would then lead to the overly long times most likely decreasing and thus the effect of not removing any of the short times should be counteracted by all the other times remaining the same, combined with the explanations above.

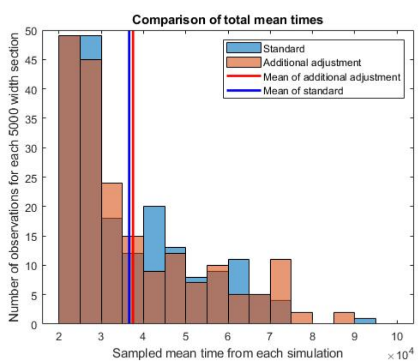
Keeping in mind the explanations above, the proposed change does show reason for being investigated further as the improvements in the model should in some manner still be replicated in reality, given enough care to which cabs are actually being sent to completion.

4.4.5 Modification 3: Additional Adjustment Station

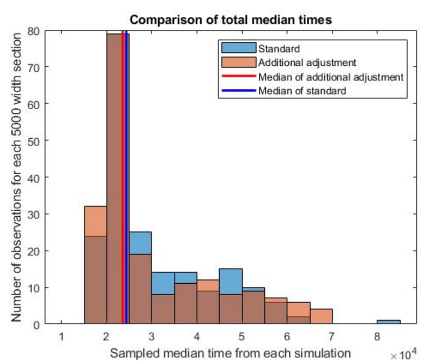
Table 15 contains information from 204 simulations of a standard week in the facility for the standard model and the model based on modification 3. The information is split into different segments of the painting process, namely Primer elevator, TCemu, Control, Completion line and the total time, with the values being of the mean and median values along with the standard deviation of these. Graphical explanations for the most important and varying periods, total, Primer elevator, TCemu and TT300 can be obtained from the histograms in Figure 24 and Figure 25, with the information in Table 15 being deemed enough to evaluate the two remaining periods.

Table 15: Comparison of simulated means, medians and standard deviations between the standard model and the model based on modification 3 with an additional adjustment station

Model	**Standard**	**Modification3**	**Standard**	**Modification3**	**Standard**	**Modification3**	**Standard**	**Modification3**
Values	Mean	Mean	*SD* of Mean	*SD* of Mean	Median	Median	*SD* of Median	*SD* of Median
Total	36607	37521	14387	16373	24338	23564	12303	13902
Primer Elevator	8539	8827	1925	2169	5482	5503	986	1197
TCemu	24588	24889	8633	9178	5725	5245	8454	9552
Control	6784	6839	1282	1382	4935	4939	643	609
TT300	10777	10852	4236	4426	2260	1945	658	619
Completion Line	2944	2805	378	348	2643	2547	545	479

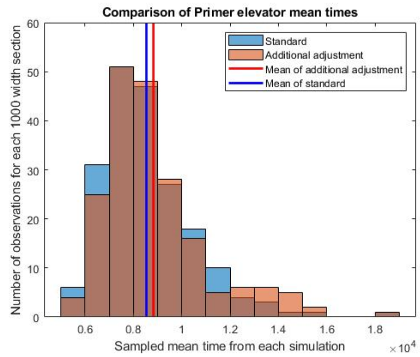


(a) Total time in the system means

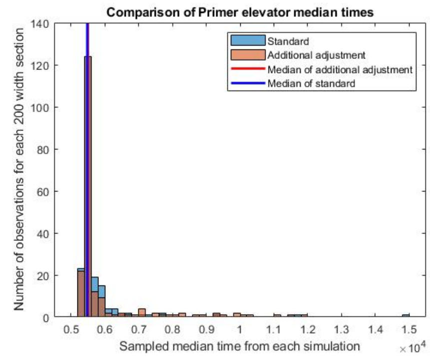


(b) Total time in the system medians

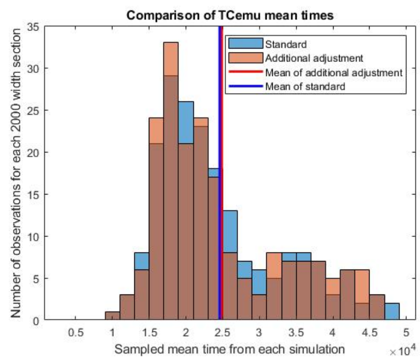
Figure 24: Comparison of mean and median times for the total process time between the standard model and the model for modification 3



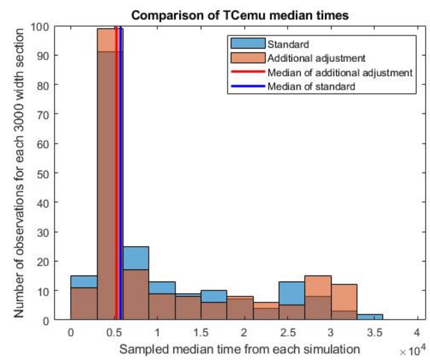
(a) Primer elevator mean



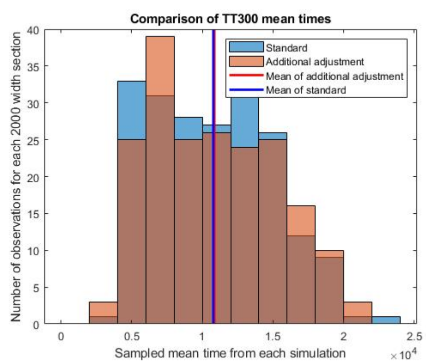
(b) Primer elevator median



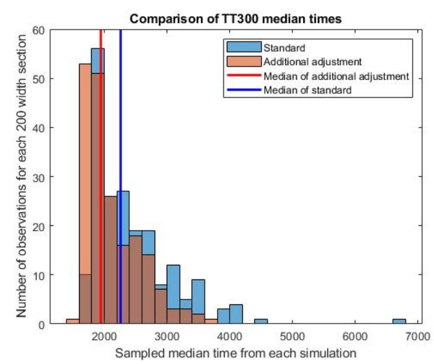
(c) TCemu mean



(d) TCemu median



(e) TT300 mean



(f) TT300 median

Figure 25: Comparison of mean and median times for the time periods of Primer elevator, TCemu and TT300 between the standard model and the model for modification 3

Given the results above it is clear that the implemented change does not have a big impact on the system at any stage of the process. It does change the weight of the values however as while the mean seems to have been moved forward, the median values have generally decreased despite an increase in the standard deviation for the total median.

4.4.6 Discussion of Modification 3

From Table 15 and Figure 24 it can be seen that while the total mean time and the variation of mean and median times have increased, the total median has been lowered. Similarly, TCemu and TT300 have both seen a slight increase in mean values, but their median values have seen big decreases. Meanwhile, the variation for TCemu has increased a fair bit while TT300 has seen a slight increase in variation of means and a slight decrease in variation of median values.

Looking at Figure 25 one can clearly see that the values around the mean and median lines for the standard model have been spaced out, with the plot for Primer Elevator means being the only slight detraction from this. While the average results do not propose that the change did anything in particular to the model, the spacing of the results does tell another story. As the biggest change has happened to TT300 medians, there is reason to look further into where this period is located and it gets affected the most.

The change implemented in the model is essentially the center point of TT300 as most cabs go through any of the adjustment sections. Because of this, the enormous influx in times below the new model's median line is understandable as more stations are available. The mean value for this period shows a different effect however. As essentially all mean values around the mean line have seen a decrease in occurrence with more extreme values occurring increasingly, the effect of more cabs coming out faster from repairs seems to have been that waiting times have become less predictable. An additional result of the change will be that more cabs can be sent on the path to the Completion Line at the same time. This could in turn mean that while cabs being sent to 0-meter level have a quicker passage to Completion Line, the ones that have not received any deviations might have a longer path as they wait for the finished cabs from 0-meter to come up into the Completion Line.

Further explanation follows what was mentioned in Section 4.4.2 about the model methodology being built to replicate the work of a process controller but without the feedback so issues can occur when the system is overfull.

Now given all the context mentioned above it is possible that the effect of an additional adjustment station is more significant than what the model makes it out to be. This, as the shortest observed median times have seen an influx in both total and TT300 times. With more feedback on how to proceed with other cabs given an increase in shorter times, there could be opportunity to capitalise on this and shorten the time spent in the process. The problem with this whole scenario is however that the time spent in the process is not the biggest issue at hand as the quantity of cabs produced is at an acceptable level. The possible shortened process, which, given the result of the model is not guaranteed, comes at the cost of two matters. First of all, in order to operate an additional adjustment station there is a need for an additional person to operate the station at all times that it is active. Given that the employed personnel are just enough to operate the facility at current time, at least three new employees have to be hired in order to gain anything from the new station which seems to be a step in the wrong direction. Secondly, the situation at Volvo does not allow for an additional station to be built due to there being no space left at the current 0-meter level. This was thus a more theoretical modification that looked to investigate how stable the system was and if anything could be altered slightly in order to obtain a different result. Seemingly that is not the case as no general change could be observed and the system can thus be seen as very stable.

5 Conclusions and Recommendations

This thesis has studied a painting process at an industrial manufacturing facility. The process has been defined from when a cab enters the painting facility to when it exits it. The purpose of this study is to use simulation in order to analyse said process. The resulting simulation model is to be able to approximately replicate the time spent in each part of the process. This, in order to indicate the effects changes in the process have and support decision making.

5.1 Review of the Study

While some parts of the inaugural data analysis were deemed surplus to requirements due to the model build, the in-depth understanding of all parts of the process proved to be very beneficial. As most authoritative personnel have one section in the process that they are familiar with, questions asked to one employee would be answered completely differently by another. At the start of this thesis, there was discussions of simply focusing on one section of the process in order to fully capture all of its complex elements. Because of the limited visibility through the process that is present throughout the facility, the decision to model the entire painting process was probably better in the end. While the model is not a perfect representation of reality, it does provide an approximate representation of how the process functions and where cabs are actually designated. The model serves its purpose of providing an overview that allows for someone without prior knowledge of the process to gather a good understanding almost instantly.

While the result of our three potential process modifications may seem pale as two of three did not alter the general state of the system at all basically, that too contains more information. Given that no slight change in the system, not even implementing a whole new adjustment station provided any great effects to the system, another connection to reality can be established. As the painting process at Volvo GTO Umeå has been built in a compact environment with a specified workload expected, slight alterations should not have a big effect on the process. If some small change in positions could be implemented to streamline the process completely then the current implementation would be flawed and as it has cost a lot of money, that should not be expected. Returning to the simulation model, slight alterations here should not provide big changes to the system either if it were to replicate reality effectively, which, is also the case as while changes happen, they are offset by changes in other parts of the system. The alteration that did prove significant was one that does not really have anything to do with the process as it consists of simply sending less cabs for repair. The result showed both a significantly faster pathway throughout but also far more stability which could reduce the need for process operators and the amount of unwanted stoppages. The realism aspect of the extent of these effects has been covered and discussed but nevertheless, it does show that the best option for improvement is to look further into whether all cabs sent for repair are actually in need of it. With a clearer framework for employees at Volvo to act upon, major improvements may be able to be brought forward without altering the process. It is thus our recommendation from the experience of this thesis that Volvo GTO Umeå redirects their efforts of improvement to gathering a more consistent framework for the top coat control station and educating their employees to follow it.

5.2 Limitations of the Study

As has been mentioned there are a multitude of simplifications that have been made in order to create the model which obviously need to be taken into account when evaluating the results. Firstly, the distinction of color groups instead of having all possible color combinations will have an impact on the color optimization buffer. In here, color sequences are planned and formed for top coat painting. The model only forms these sequences based on color group which could affect the time there and in the top coat painting. All of the complexity at 0-meter level has not been captured either as there is no distinguishing between what deviation has been found or where it is. Additionally, buffers have been combined at times in order to counteract Simulink's randomness in how the picking system works. The same systematic also causes excessive outliers at times which could and probably should be ignored when looking at individual and average effects of changes as these values can be skewed heavily.

The most obvious limitation is one that has already been brought up multiple times, the absence of process controllers. As they have the purpose of making real time interactions with the process in order for long stoppages to be avoided, for cabs to be prioritised if they have been in the system for a long time and to correctly utilise all available buffers, issues related to these can happen in the model. The distinction between more shorter times and more way longer times than the median of data is most probably caused by this and does need to be taken into consideration when evaluating the result.

References

Asbjörnsson, G., Hulthén, E. and Evertsson, M. (2013), 'Modelling and simulation of dynamic crushing plant behavior with matlab/simulink', *Minerals Engineering* **43**, 112–120.

Banks, C. and Sokolowski, J. (2010), *Modeling and Simulation Fundamentals: Theoretical Underpinnings and Practical Domains*, John Wiley & Sons.

Björch, T. and Strålberg, F. (2016), 'Customer support process analysis: Using statistics and modeling to analyze a global customer support process'.

Blagotić, A. and Daróczi, G. (2022), *Rapport: a report templating system.* R package version 1.1.
URL: *https://cran.r-project.org/package=rapport*

Chihara, L. M. and Hesterberg, T. C. (2011), *Mathematical statistics with resampling and R*, 1 edn, John Wiley & Sons.

Davison, A. C. and Hinkley, D. V. (1997), *Boostrap methods and their applications*, Cambridge University Press, New York.

Law, A. M. (2015), *Simulation Modeling & Analysis*, 5 edn, McGraw-Hill.

MATLAB (2022), *version 9.12.0 (R2022a)*, The MathWorks Inc., Natick, Massachusetts.

Peacock, S. (2002), The use of simulink for process modelling in the sugar industry, *in* 'Proc. S. Afr. Sug. Technol. Ass', Vol. 76, pp. 444–455.

R Core Team (2021), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Ross, S. M. (2013), *Simulation*, 5 edn, Academic Press.

Sargent, R. G. (2011), Verification and validation of simulation models, *in* 'Proceedings of the Winter Simulation Conference', WSC '11, Winter Simulation Conference, p. 183–198.

A Appendix - Pathway probabilities for top coat 0 meters level

Table 16: The probabilities of where the cab will be destinated next given the station at top coat 0-meter level it has been at and which color group it belongs to

Position of the Cab	Color Group	OK	Adjustment	Improvement	Panel Improvement	Grinding	Masking	Truck Table	COB
Adjustment	Normal	90.59%	N/A	6.54%	1.30%	1.10%	0.06%	0.36%	0.06%
Adjustment	Uncommon	83.67%	N/A	11.94%	2.19%	1.16%	0.32%	0.58%	0.13%
Adjustment	Metallic	87.7%	N/A	6.04%	2.70%	3.18%	0.16%	0.16%	N/A
Improvement	Normal	71.10%	27.11%	N/A	N/A	0.16%	1.62%	N/A	N/A
Improvement	Uncommon & Metallic	62.12%	34.70%	N/A	N/A	1.06%	2.12%	N/A	N/A
Panel Improvement	Normal	75.00%	22.86%	N/A	N/A	0.71%	1.43%	N/A	N/A
Panel Improvement	Uncommon & Metallic	64.17%	34.22%	N/A	N/A	0.53%	1.07%	N/A	N/A
Grinding	All Colors	N/A	N/A	1.07%	N/A	N/A	7.58%	23.23%	68.18%
Masking	All Colors	N/A	N/A	57.14%	42.86%	N/A	N/A	N/A	N/A