MSc thesis in Civil Engineering for Transport & Planning

Verification of Cyclist Simulation Model

C.-Y. Chou

2020

VERIFICATION OF CYCLIST SIMULATION MODEL

A thesis submitted to the Delft University of Technology in partial fulfillment
of the requirements for the degree of

Master of Science in Civil Engineering for Transport & Planning

by

Chen-Yen Chou

2020

The work in this thesis was made in the:

Active Mode Lab
Department of Transport & Planning
Faculty of Civil Engineering
Delft University of Technology

Student Number :	4900278
Committees :	Dr. ir. Winnie Daamen
	ir. Alexandra Gavriilidou
	Dr. ir. Yufei Yuan
	Dr. ir Simeon Calvert

PREFACE

As part the curriculum of the Msc programme in Transport & Planning at TU Delft, the thesis work is considered a final achievement of the programme. The journey cannot be finished without the support from other people. First of all, I want to thank my supervisor Dr. ir. Winnie Daamen for giving me this project to explore the fascinating world of the cyclist simulation. I would also like to show my gratitude to Dr. ir. Yufei Yuan and Dr. ir. Simeon Calvert for sharing their expertise as well as their critical advice on the project. Additionally, I am grateful to my daily supervisor ir. Alexandra Gavriilidou who started the thesis work with me. Throughout every biweekly meeting, they provide me with the critical thinking and direct me through the whole research. Last but not least, I would like to thank my friends and family for their continuous support, especially during the pandemic.

SUMMARY

Model verification is a crucial step in the development of a simulation model. It determines the quality of a simulation model when transforming a conceptual model into a simulation model. The correctness of the logic and the theory implemented in a simulation model is ensured through model verification.

The studies on the methodologies to verify a vehicular traffic simulation model have been developed for a long time, but limited amount of research was found related to the verification of cyclist simulation model. A simple way to verify a cyclist simulation model could be applying those methodologies used for vehicular simulation model. However, with a review of the existing literature, it is known that there was no standard procedure to conduct model verification. With the intention to fill up this gap, this thesis is established to create a framework for model verification based on the past methodologies used to verify traffic simulation models. To reach this goal, a research question has been formulated:

What is the generic framework to verify traffic simulation models with the consideration of cyclists that hitherto has not been discussed?

To create such a framework, the definition and the purpose of model verification are derived from the literature related to the development of simulation model. In this research, model verification is defined as ″a process to justify the implementation of a simulation model by comparing the simulation model with the conceptual model using system data″. Two objectives of model verification are recognised: (1) ensure the basic functionality of the simulation model by reproducing the result of the conceptual model which the simulation model is based on or by performing the intended application of the simulation model; (2) understand the input-output relation in the simulation model in order to determine the applicable range of the parameters used in the simulation model. These two objectives are used to rank the level of model verification, level 1 and level 2, respectively and level 1 should come prior than level 2. The essential procedures of model verification are then derived from the studies related to the verification of traffic simulation models.

The use of the framework is intended to make a structural and traceable approach to verify a model such that the framework can be applied to any traffic simulation model. To incorporate the existence of cyclists simulation models with the framework, a research on the microscopic cyclist simulation models is also made to explore the potential difference between the verification of vehicular simulation models and the verification of cyclist simulation models.

According to the findings from the literature review, the framework for model verification is created. It contains a start procedure and seven main steps, which are categorised into the two phases of verification, ″Creation of Plan″ and ″Execution of Plan″. The first phase, ″Creation of Plan″, includes step ″Determine level of verification″, ″Module Identification″, and ″Verification planning″. The second phase ″Execution of Plan″ contains step ″Perform tests″, ″Analyse problem″, ″Modify simulation model″ and ″Modify plan″. Each of the steps consists of supportive sub-steps or subjects to accomplish the main step. The framework is presented in Figure I.

The procedure starts with questions to ensure the framework is applied with a traffic simulation model that requires a practice of verification in its development.

In phase 1, the steps are performed to collect the information from the input simulation model so that a suitable plan can be created to verify the input model. After that, steps in phase 2 are performed to execute the plan with the consideration of the possible issues that could occur in the plan or in the input model.

The framework is constructed through a thorough exploration of the literature related to the verification of vehicular simulation model. Because of that, the applicability of the framework to the simulation models other than vehicular simulation model remains a question. To test whether the framework can be applied to other types of simulation model, such as a cyclist simulation model, a case study is performed. With the feasibility as a cyclist simulation model and the availability of the model for the author, the queue formation simulation model is selected to best assess the framework at the same time complete its verification.

The assessment of the framework is done by applying the framework to the cyclist queue formation simulation model. In phase 1, the information about the cyclist queue formation simulation model is collected. According to the information, the purpose of the verification is to ensure the model's basic functionality, and there are seven modules in the simulation model. Based on that, a plan for level 1 of verification to achieve the purpose is created. The plan was then executed in the phase 2 of the framework. During the execution, one of the modules, "DCM Physical", was not able to pass the test indicated in the plan due to the inappropriate selection of the verification technique later found out with step "Analyse problem" performed. The problem was related to the plan itself and thus step "Modify plan" was performed to fix the problem. After that, the modified plan was completely performed and therefore the application of the framework was accomplished.

From the application of the framework, it is known that all of the required steps in the framework can be performed, and one of the steps, "Modify simulation model", is not necessary in this case. Also, the queue formation simulation model proved its capability to fulfill its aiming objective of model verification with the framework. Based on these results, it can be concluded that the applicability of the framework is ensured with the case study.

The pros and cons of the framework are summarised from the application of the framework. The main advantage of the framework is its systematic structure in verification. The benefits of the systematic structure is the easiness to follow, a traceable process of verification, and comparable process and product of model verification. There is no structural disadvantage found in the framework, but the requirement of the conceptual model may limit the application of the framework. To conclude, the framework provides a convenient and reliable measure to conduct model verification.

Besides, room for improvement in the queue formation simulation model was indicated during the verification. A potential problem which could affect the later development of the model was found. It is possibly related to the parameters used in the "DCM Physical" module. In the verification, "DCM Physical" displayed its intention to describe the behavior of cyclists when approaching the stop line at an intersection. The correction of steering and pedalling behaviour of the cyclists in the simulation model was found reacting slower than the cyclists in the real world, resulting in considerable deviations of the final stop positions of the cyclists in each run of simulation. This could be caused by the rather long update interval currently implemented in the model. A potential solution is to lower down the update interval. However, this solution requires the understanding of the influence in parameters used in the model, which leads to a necessity of a verification of level 2.

To conclude, this research has created a framework and proved its applicability to verify a cyclist traffic simulation model. In terms of the products of the research, the framework is built up from the literature with a brief discussion on the microscopic cyclist simulation models. The framework showed its value with the case study. Its organised structure provides a systematic measure for model verification. Also, the queue formation model is verified within the study area of its conceptual model after completing a verification of level 1 using the framework. The room for improvement was also indicated for the further development of the model.

However, there are some points should be discussed. First, the planning for verification of different level is only briefly discussed. The selection of a statistical test in level 1 and the design of a sensitivity analysis in level 2 are both worth elaborating more. Second, the possibility to verify a model by comparing two models is not discussed in the framework though it is found work for model validation in the literature. Third, the framework of this research is only applied and tested with a single case at the queue formation simulation model. This makes the credibility of the framework rather limited. By increasing the number of cases, it is possible to further justify the applicability of the framework to other traffic simulation models. Last, the indication of the room for improvement of the queue formation simulation model brought the necessity of a verification of level 2 for the model.

Therefore, in terms of future research, it is recommended to: i) expand the framework with more detail in the planning of verification; ii) include the possibility to use comparisons between different models as a verification measure in the framework; iii) conduct justification of the framework on multiple cases, which should not be limited to vehicular or cyclist simulation model only, but also other traffic simulation models; and iv) further verify the queue formation simulation model in a verification of level 2, where the input-output relation and the applicable range of parameters are to be explored.

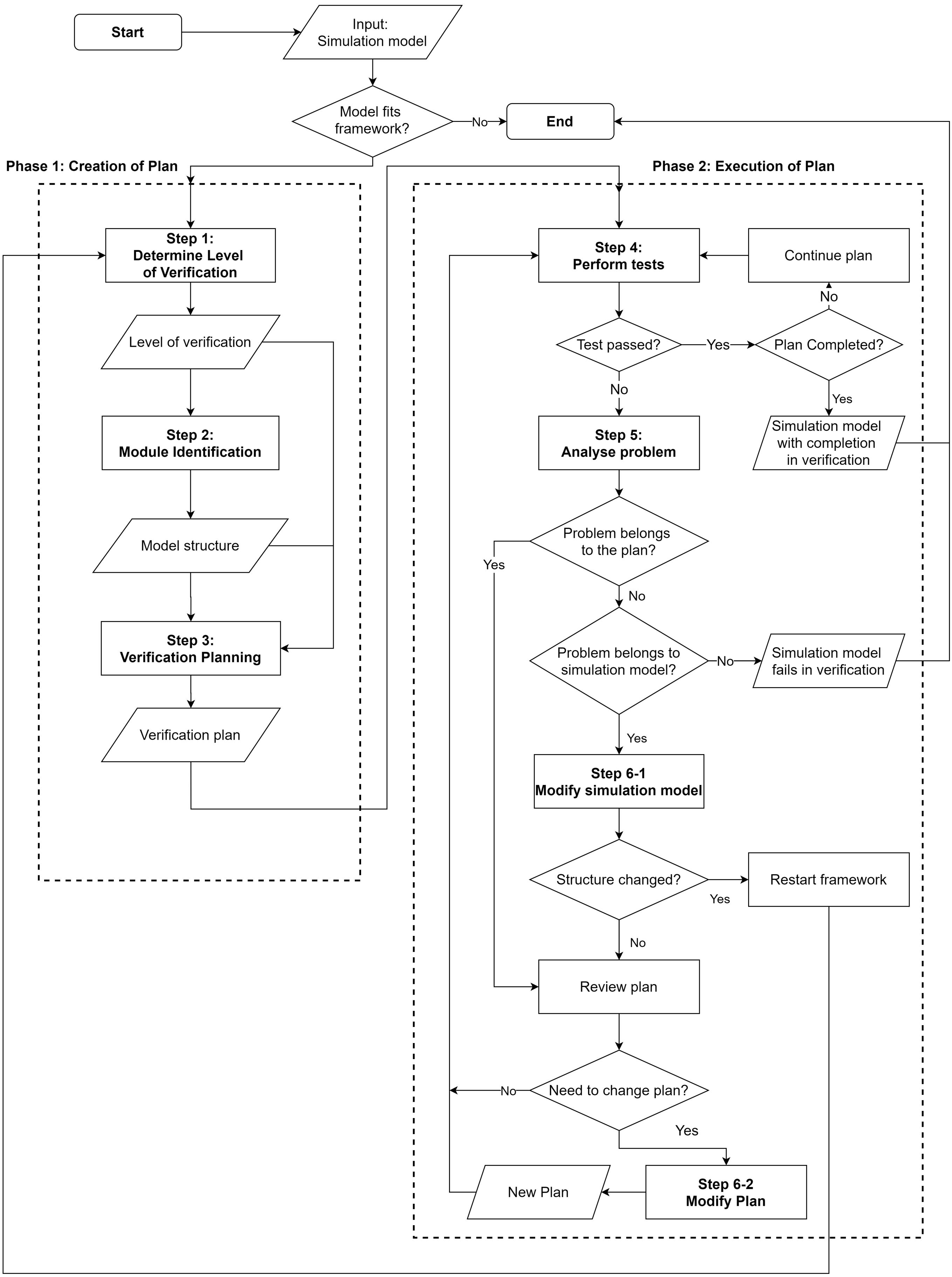

Figure I: Developed framework for traffic simulation model verification

CONTENTS

1 INTRODUCTION

With the growing of the population and the rise of the transport network companies such as Uber, Lyft, and Bolt in the urban area, many cities are suffering from the increase of traffic congestion, and it will be more serious in the future if the transportation system remains the same. As the competition between different modes for the limited urban space reserved to transport is growing with urbanization, there is a need to manage this space in order to improve accessibility for everyone [Gonzales et al., 2010]. Previous studies have shown that bicycles and public transport are much more efficient than cars in terms of use of space [Dekoster and Schollaert, 2000; Botma and Papendrecht, 1991]. Walking and cycling then complement the new mobility framework by providing a convenient and sustainable option of individual mobility [Lowe, 1990]. Accordingly, increasing the multi-modality of the system could be a potential solution.

Cycling as a mode of transportation has many advantages for both cyclists and society: it is a low-cost, low-polluting, health-improving way to travel, and a growing number of cities throughout the world are implementing policies to promote cycling [Handy et al., 2014]. Perhaps as a result of that, bicycle use had increased 6 % from 2017 to 2018 around the world [Eco-Counter, 2019]. In order to intrigue more people to use a bike, it is essential to assess the current design of the cycling infrastructure to guarantee the efficiency and safety for the users. A simulation of cyclists with its quality ensured by verification, validation and calibration is helpful to assess different designs of infrastructure and traffic management in order to point out the problems. However, as Gavriilidou et al. [2019a] said, "even when the interest in cycling in cities increases, research on bicycle traffic behaviour is still in its infancy", not to mention the cyclist simulation studies.

The Active Mode Lab (under Department of Transport & Planning, Faculty of Civil Engineering and Geosciences) in TU Delft is currently developing a cyclist queue formation simulation model based on a conceptual model developed by Gavriilidou et al. [2019a]. The queue formation model is a microscopic operational behaviour model, using a discrete choice model to describe the cyclist's decision-making process. The model has been implemented in Matlab to test its feasibility, and a verification process remains to be done.

Model verification is a critical process in developing a traffic simulation model [Horiguchi and Kuwahara, 2002]. Although the definition of "model verification" is sometimes mixed up with "model validation" or "model calibration" [Rakha et al., 1996], the aim to carry out a model verification in different studies is the same, which is to justify the development of a simulation model by a certain measure. There are many studies regarding the definition, the process or the application of model verification and validation on vehicular traffic simulation model [Benekohal, 1991; Rakha et al., 1996; Horiguchi and Kuwahara, 2002; Sargent, 2010]. Nevertheless, verification of cyclist traffic simulation model has never been mentioned in the previous studies. A simple way to verify a cyclist simulation model could be applying the methodologies used in the verification of vehicular simulation models. However, due to the fact that simulation models are developed for different purposes, the verification process may vary to for different models and thus no standard procedure could be found in the literature.

To fill up this gap, this thesis attempts to create a generic framework as a standard procedure for traffic simulation model verification with the specific consideration of the cyclist simulation models.

In this chapter, the problem analysis is made with the problem statement and the research objectives. The research questions are then formulated according to the objectives of this research. Following that, the research approach is introduced, along with the research steps. The contributions of this research are addressed next. After that, the structure of the report is briefly explained.

1.1 PROBLEM ANALYSIS

In this section, the problem statement is made based on the gaps found in the coarse search of literature. The according research objectives are addressed based on the problem statement.

1.1.1 Problem Statement

With a search of the existing literature, it is found that there are many studies discussing the verification and validation of vehicular simulation model. However, the process to conduct model verification differs with the purpose to develop the model. There is not a standard procedure to follow when conducting model verification. Therefore, it is necessary to establish a framework representing a standard procedure to conduct model verification for traffic simulation.

The framework is to be created according to the literature related to the verification of traffic simulation models, where most of the studies are relevant to vehicular simulation model. To understand whether the framework can be applied to verify other types of simulation model, a case study is required.

Considering the fact that there is not really a research focusing on the verification of cyclist simulation model, applying the framework to a cyclist simulation model and discussing the difference to verify a cyclist simulation model becomes credible for the future studies of cyclist simulation.

As the Active Mode Lab in TU Delft is implementing the cyclist queue formation model into a simulation environment and the next step is to conduct a model verification, it is plausible to use this cyclist simulation model as a case study to find out this problem for the availability and the feasibility of the model.

1.1.2 Research Objective

Based on the problem statement, the objectives of the research is addressed:

1. Create a standard verification procedure, as a framework, for model verification of traffic simulation models.
2. Test the applicability of the framework on verifying a cyclist simulation model with the application of the framework to the cyclist queue formation simulation model.
3. Verify the cyclist queue formation simulation model with suggestions to improve the current model.

4. Discuss the potential differences between verifying a cyclist simulation model and verifying a vehicular simulation model.

1.2 RESEARCH QUESTIONS

To achieve the research objectives, the main research question is formulated:

What is the generic framework to verify traffic simulation models with the consideration of cyclists that hitherto has not been discussed?

To find out the answer to the main question, the following sub-questions are formulated with a few supporting questions as assistance:

Q1-1 *What is model verification for traffic simulation model?*

Before everything starts, it is necessary to totally understand what "model verification" means, especially when it is found in a short literature study that the definition of model verification varies with the studies. To answer this question, the following supporting questions are made:

Q1-1.1 *What is the definition of model verification?*

Q1-1.2 *What is the purpose to perform model verification?*

After the meaning of model verification is understood, it is possible to go for the objectives of this research.

Q1-2 *What methodologies that were used to verify vehicular traffic simulation model are found in literature?*

To create a framework, it is important to perform a literature review on what methodologies were used. To answer this question, the following supporting questions are made:

Q1-2.1 *What was the process to perform model verification?*

Q1-2.2 *What were the approaches applied in model verification?*

Q1-2.3 *How were the process and approaches of model verification determined?*

The answer to the question should be found in the literature of the methodologies, process, procedures, approaches or application that were used to verify traffic simulation model.

To understand the risk to apply the framework to a cyclist simulation model, the following question is asked:

Q1-3 *What are the potential differences to verify a cyclist simulation model from verifying vehicular simulation model?*

With the aforementioned questions answered, the results of the literature study should provide adequate information that created a theoretical base for the framework, resulting in the next question:

Q1-4 *According to the methodologies found in literature, what are the steps to be included in the structure of the framework for model verification?*

A framework should be set up based on the literature. The framework includes certain procedures and core activities. After the framework is created, the next question is formulated to assess the framework with an application:

Q1-5 *How is the framework performed in the application?*

To assess the proposed framework, a case study is required. As aforementioned, the cyclist queue formation simulation model is selected for its availability and feasibility. The applicability of the framework could be discussed through the case study. Besides the performance of the framework, the pros and cons of the framework should also be addressed. This leads to another question:

Q1-6 *What are the pros and cons of the framework that discovered in the application?*

With all steps introduced and performed, the pros and cons of the framework can be discussed. These pros and cons are supported by the performance from the case of the queue formation model, and they would leave room for the future research.

Another objective of this research is to provide suggestion to the current model, so it brings to the last question:

Q1-7 *If there is room for improvement of the simulation model, what is it?*

No matter how the framework would perform, the identification of the potential problems in the model is important for the development of the simulation model.

With the above sub-questions answer, it should be possible to answer the main research question.

To find the answers to all the sub-questions and the main research question, this research is constructed in four parts: (1) studying the existing verification methodology, (2) creating the framework, (3) applying the framework on a case and (4) evaluating the application and the framework. The detail of the research approach is discussed in the following section.

1.3 RESEARCH APPROACH

To deal with the sequence of research questions addressed in the previous section, a suitable research approach is required. This approach is mainly supported by a literature review and a case study. The structure containing the research steps of this study is illustrated in figure 1.1.

This research starts with a literature review. Information regarding model verification of traffic simulation model and the potential difference to verify a cyclist simulation model is to be collected in the literature review. The results of the literature review are used to create a framework with according steps derived from the past methodologies to verify traffic simulation models.

After the framework is created, a case study to confirm its applicability is then conducted, where the framework is applied to verify a cyclist simulation model. The framework is assessed with the results of the application. According to Leutscher [2018], a newly developed methodology can be assessed with two indicators: 1) the ability to perform all the actions and 2) the evaluation of the product. Therefore, the assessment of the framework includes two parts, the ability to perform the framework and the evaluation of the tested simulation model. The pros and cons to use the framework are addressed based on the experience learned in the application.

At the end, there will be three products when the research is finished. One is the development and the demonstration of a framework that can be used to verify traffic simulation model. One is the queue formation simulation model completed in verification. The other one is the suggestion of the queue formation simulation model for the further development of the model.

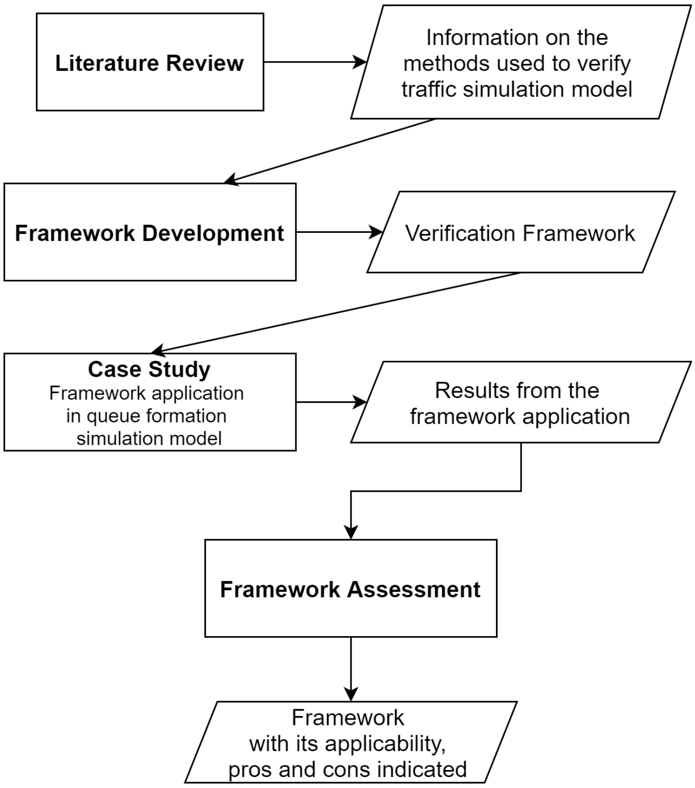

Figure 1.1: Research Approach

1.4 RESEARCH CONTRIBUTIONS

The research contribution includes two parts: scientific contribution and practical contribution.

Scientific contribution

In the filed of simulation studies, this research aims to address the challenge of developing a generic verification framework for traffic simulation model, including the consideration of cyclist simulation model. This framework can be used to verify any types of simulation models in a systematic way, where the process and the product of model verification become more comparable with the use of the framework. Besides, by carrying out a case study of the framework on a cyclist simulation model, the understanding of the verification procedure for the cyclist model could be helpful for developing new cyclist simulation models.

Practical contribution

Another contribution of the research is systematically verifying the queue formation simulation model with adequate information to improve the model so that the Active Mode Lab in TU Delft can further validate and calibrate it, applying it as a valid simulation tool to assess any potential traffic management for bike traffic, even for mixed traffic. Therefore, the practical contribution of this study is in the form of a verified queue formation model that brings out the potential to assess the infrastructure design from another perspective with higher accuracy.

1.5 REPORT STRUCTURE

The paper is structured as follows. Chapter 2 is "Literature Review", focusing on the findings from the literature related to the verification of traffic simulation model and the state-of-art cyclist simulation models. An overview of the queue formation model is also included in this chapter. In chapter 3, "Framework Development", the explanation of each step of the framework created based on the findings from the literature review is addressed. Chapter 4, "Framework Application", describes the application of the framework to the cyclist queue formation simulation model. Following that, the assessment of the application is addressed in Chapter 5, "Assessment of framework application". Last, the research findings along with the discussion and the recommendation are discussed in Chapter 6, "Conclusions, Discussions and Recommendations".

2 LITERATURE REVIEW

As the research objective, questions, approach, contribution and the structure of this thesis are presented, we can move to the first part of the research content: literature review. This chapter is dedicated to deal with previous research on traffic simulation verification along with the state-of-art cyclist simulation models. The objective is to explore the methodologies that were used to verify traffic simulation model and create a potential framework to verify simulation models, including cyclist models. The focus is on the methodologies to verify a traffic simulation model. The procedures, along with the techniques that were applied to verify a model, especially the decision-making process to choose certain techniques are expected to be the core actions for this framework. Aside from that, the difference between verifying vehicular simulation model and verifying cyclist simulation model is to be discussed with the literature related to cyclist simulation models.

There are three subjects in the literature study: 1) development of traffic simulation model, 2) model verification of traffic simulation model and 3) cyclist simulation models. Within the first subject, the development of traffic simulation model is discussed. The meaning of model verification in a simulation model development is explained thoroughly with a clear definition.

Next, in the second subject, the difference between model verification and model validation is illustrated. The process of model verification is explained with examples from the previous studies found in the literature.

Last, in the third subject, the literature study focuses on the microscopic cyclist simulation models, where the difference in the individual behaviour is described explicitly. The characteristics of different microscopic cyclist simulation models is addressed to expose the potential difference to verify a cyclist simulation model. An overview of the cyclist queue formation model [Gavriilidou et al., 2019b] is also introduced.

Finally, the essential procedure to be included in the framework is discussed in section 2.4. The information gained in the literature study is analysed and summarised, thus the results of the literature study are obtained.

2.1 DEVELOPMENT OF TRAFFIC SIMULATION MODEL

Traffic simulation is used to represent a complex system in the real world, such as a traffic environment, as a more manageable system with observable and controllable parameters. This manageable system is called a ”model”. By translating the model into a simulation environment with a programming language that a computer can understand, a traffic simulation model can be obtained. A traffic simulation model should reproduce the behaviour of the system, such as the traffic on the road. The output of a traffic simulation model is the accumulated statistics which are used to analyse the traffic behaviour. A good traffic simulation model, which is able to precisely represent the system being modeled, can be used as an analytical tool to assess the current transportation system. The results of a simulation can be used as a basis for selecting the best traffic operation alternatives. Also, a reliable traffic simulation model can be used as a research tool to understand the behaviour of traffic flow under different scenarios.

Generally, traffic simulation models are classified as macroscopic, microscopic and mesoscopic model [Kessels, 2019]. In the macroscopic models, traffic behaviour is represented in terms of aggregate measures such as traffic volume, space mean speed, and average density. Macroscopic simulation models are designed to deal with large scale problems. These models do not show the detail of individual behaviour in traffic, but they are efficient.

In the microscopic models, the individual traffic participant is represented by a set of variables such as speed, position, acceleration, etc., updated at constant time steps. These models are highly detailed and more precise in terms of individual behaviour. However, they require larger computing time and their accuracy in aggregate measures need to be calibrated [Helbing et al., 2002].

Mesoscopic models were developed to fill the gap between the microscopic models the macroscopic models. They describe traffic flow in aggregate terms, for example probability distributions, with the behavioral rules for road users defined individually. In addition, there are hybrid mesoscopic models that appeared recently as the combination of microscopic and macroscopic models [Kessels, 2019].

The development of a simulation model is not easy work. There are many processes to go through to reach a valid simulation model. As the focus of this thesis is model verification, the position of model verification in the development is to be found out in this literature review. Also, since the difference between vehicular simulation model and cyclist simulation model is more easily to be observed in microscopic level and the queue formation simulation model is a microscopic model, the discussion of the difference between verifying vehicular simulation model from verifying cyclist simulation model will be addressed based on microscopic cyclist simulation model.

In this section, an overview of model development is illustrated first. The definition and the purpose of model verification in the development of traffic simulation model are then discussed.

2.1.1 Model development process

In the study of Balci [1998], he presented a life cycle of a simulation study and recognised the steps that needs verification, validation and testing (VV&T) with a case study as an example. The life cycle is shown in figure 2.1. In the first three steps (Problem Formulation, Investigation of Solution Techniques and System Investigation), the problem and the system that contains the problem in the real world is identified. The potential solution to the problem is investigated and proposed. The system and the objective of the simulation are therefore defined. In the following three steps (Model Formulation, Model Representation and Programming), the objective and system is formulated as a model. The conceptual model first appears in the mind of the model developer and represented as a communicative model for people to understand. By implementing the model with a programming language, a programmed model is made. In the last three steps (Design of Experiments. Experimentation, and Presentation of Simulation Results), a plan to gather the desired information with affordable cost is made, executed. An experimental model is created from the programmed model incorporating with a formulated plan. An experiment is done with the experimental model. The results are interpreted to the stakeholders/decision-makers who are interested in solving the problem. The process of redefinition is carried out to update the model for improvement or modification.

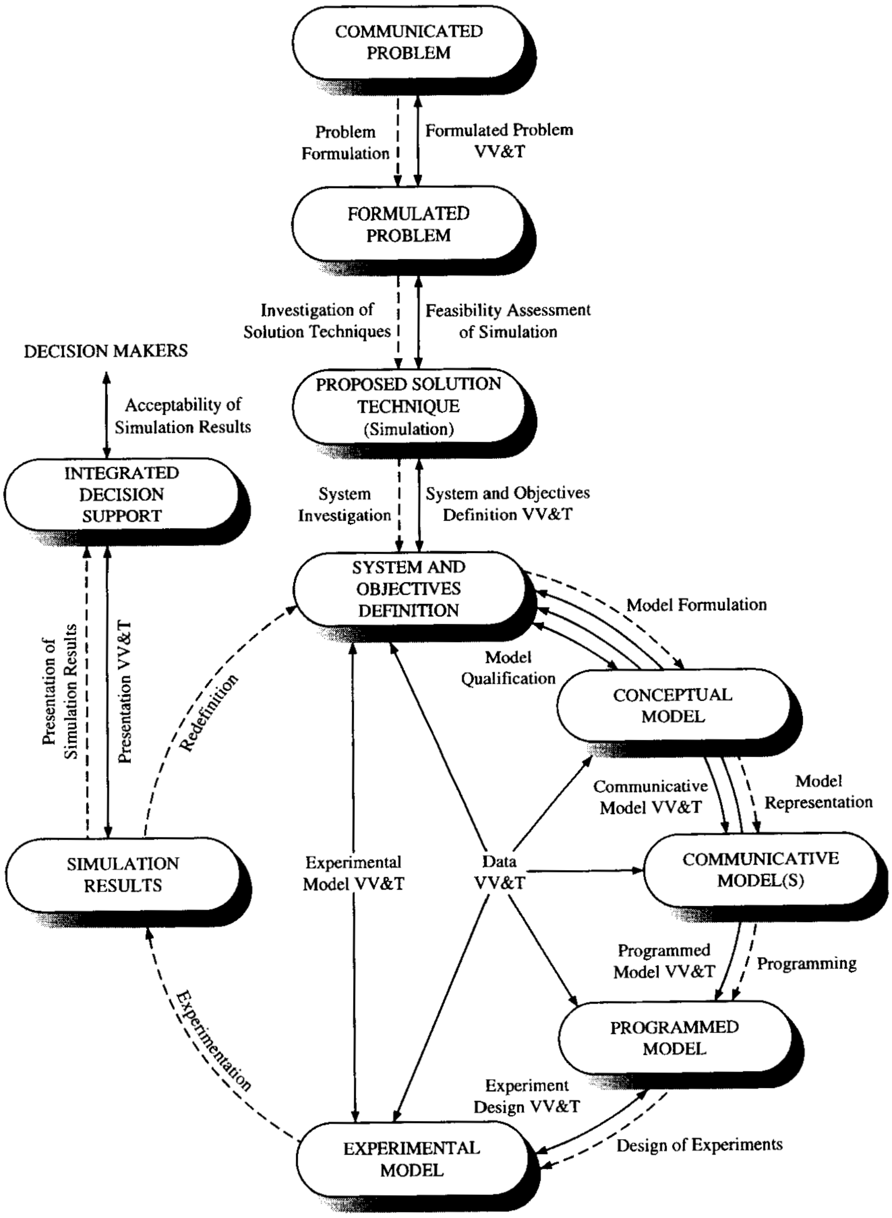

Figure 2.1: Life Cycle of Simulation Study from Balci [1998]

In another research, Sargent [2010] presented the development of a traffic simulation model in another way. He classify the steps of simulation model development into the ones that belong to the real world and the ones that belong to the simulation world. The connection between the two worlds is the system theories, which describe the characteristics and the behaviour of the system. The development process also starts with identifying the system experiment objectives. All the relative steps are shown in figure 2.2.

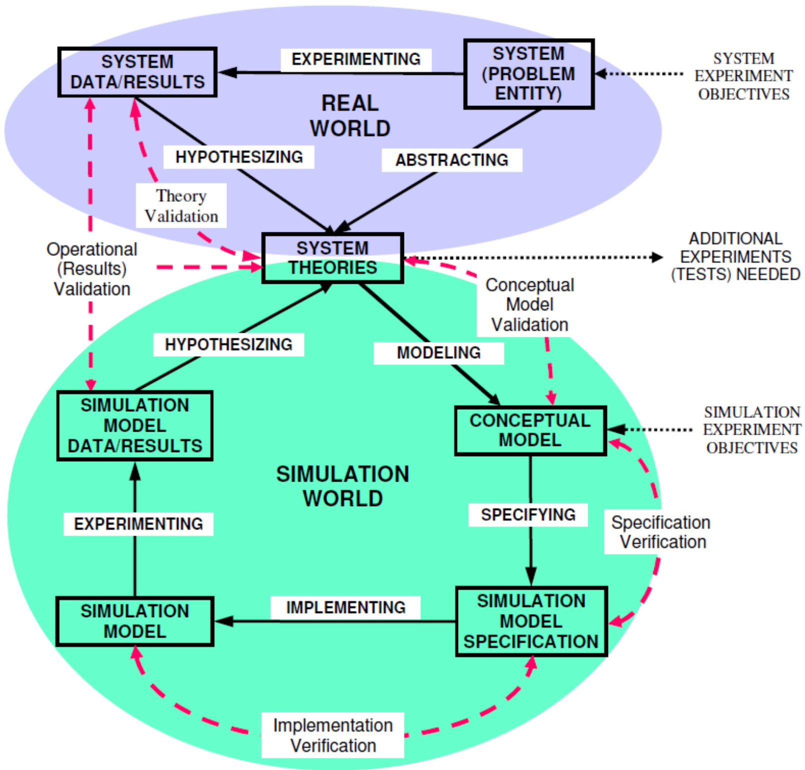

Figure 2.2: Real-World and Simulation World Relationship [Sargent, 2010]

According to Sargent [2010], to build up a traffic simulation model, a system(problem entity) is first abstracting as system theories (System Abstracting), entering the simulation world from the real world. A conceptual model, representing the problem entity with mathematical/logical representation developed for a particular study, is made through System Theories Modeling. A simulation model is then constructed from the specification (Conceptual Model Specifying) and implementation (Conceptual Model Implementing) of the conceptual model. Specification means a detailed description of the software design for programming the conceptual model, while implementation means the execution of the specification that makes the conceptual to run in a particular computer system (simulation environment), where experiments can be conducted. The outputs of the simulation model are obtained through experiment (Simulation Model Experimenting).

In conclusion, Balci [1998] provided a detailed description of the steps in the development of traffic simulation model. On the other hand, Sargent [2010] introduced a concise process with the identification between real-world and simulation world. Both studies describe a similar process to develop a traffic simulation model, but in different ways. From the two studies, the author concludes the three essential stages and the two main processes in the development of simulation model (figure 2.3).

The three stages are Problem Entity, Conceptual Model, and Simulation Model. They are categorised in table 2.2. Problem Entity is the system to be modelled. It could be an idea, a situation, a policy, or phenomena in the real world or in a proposed virtual world. It covers the definition of the system (problem entity) from Sargent [2010] and the Communicated Problem, Formulated Problem, and System and Objective Definition from Balci [1998]. Conceptual Model uses the same definition used in the study of Sargent [2010], a mathematical, logical, or verbal representation of the problem entity developed for a particular study. It also covers the definition

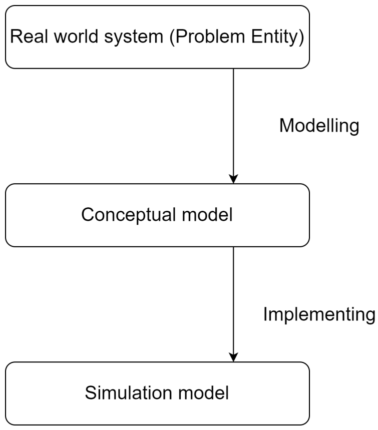

Figure 2.3: Three stages and two processes in development of simulation model

of the conceptual model and the communicative model from Balci [1998]. Lastly, Simulation Model is the computerised version of the conceptual model that implemented in a computer system and ready for simulation experiment. It comprises the definition of the programmed model and experimental model from Balci [1998] and the simulation model specification and simulation model from Sargent [2010]. In addition, the problem entity is the only stage that is in the real world. The other two stages, conceptual model and simulation model, are the stages that are in the simulation world as identified by Sargent [2010].

Table 2.1: Three stages of Simulation Model Development categorised by the author

	Sargent (2010)	***Balci (1998)***
Problem Entity (Real World)	System (Problem Entity)	Communicated Problem Formulated Problem System and Objective Definition
Conceptual Model (Simulation World)	Conceptual Model	Conceptual Model Communicative Model
Simulation Model (Simulation World)	Simulation Model Specification Simulation Model	Programmed Model Experimental Model

The two processes proceeding from one stage to another stage in the development of traffic simulation model are Modelling and Implementation. They are categorised in table 2.2. Modelling includes Problem Formulation [Balci, 1998], Abstracting of the problem, Modeling of the system[Sargent, 2010], Investigation of Solution Techniques, System Investigation, Model Formulation, and Model Representation [Balci, 1998]. Implementing describe the process of programming the conceptual model into a simulation environment, which includes the steps Programming [Balci, 1998], Specifying and Implementing [Sargent, 2010] mentioned above.

Table 2.2: Two processes of Simulation Model Development categorised by the author

	Sargent (2010)	*Balci (1998)*
Modelling	Abstracting Modelling	Problem Formulation Investigation of Solution Techniques System Investigation Model Formulation Model Representation
Implementing	Specifying Implementing	Programming

However, the development of a traffic simulation model is not a one-way procedure. Instead, it is an iterative process with assessment and justification of each step, as the bidirectional arrows displayed in both figure 2.1 and figure 2.2 from the two studies. The most often seen words to present these justifications in the mentioned studies are "validation" and "verification", and they are to be discussed in detail in the next subsection.

The author has no intention to ignore any of the steps to develop a simulation model. Instead, the more detailed the steps are described, the higher the quality of a model can be expected. Still, the author believes that the concluded version of the process gives a more intuitive aspect to understand the development of a traffic simulation model.

In this subsection, the three essential stages of developing a traffic simulation model are recognised with the process from one stage to another stage. Nevertheless, the justification of each process is not discussed yet. There are two types of justification in the development of simulation model, which are model verification and model validation. The definitions and detailed explanations of the two types of justification are exploited in the next subsection.

2.1.2 Justification in model development

Model verification and validation are two important steps of justification in the development of a simulation model. However, according to Rakha et al. [1996], "there appears to be little uniformity in the definition of each of these processes". In their study, they defined a nomenclature for these, where *model verification* is considered as the process to determine whether the computer code capture the logic of the model correctly as specified by the model designer; and *model validation* is defined as the process to determine to what extent the mechanism of a model represent the problem entity, as demonstrated by field data.

On the other hand, in the study of Horiguchi and Kuwahara [2002], model verification is defined as a process consisted of several examination tests to confirm fundamental functions of the simulation model. In another research, Benekohal [1991] used the definition of model verification as to check if the computer simulation model behaves as assumed by the experimenter, and the definition of model validation as to test whether the simulation model represents the system in the real world reasonably. Also, Kleijnen [1995] gave a more straightforward definition for both processes. He explained that model verification is to determine a simulation model that performs well by debugging the computer program, and model validation is to determine whether the conceptual model represents the system under study correctly.

From the studies mentioned above, it can be noticed that model verification is more often to be deemed as a process to check the correctness of the computer program within the simulation model [Rakha et al., 1996; Kleijnen, 1995; Benekohal, 1991], while model validation is clearly defined as a process to determine whether the simulation model approximates the system in the real world appropriately by most of the studies.

However, sometimes the process to perform model verification and validation are mixed, which might be confusing. In the research of Balci [1998], verification, validation and testing (VV&T) are treated as a combination of measure to justify whether a certain step is practised correctly. Because of that, there are many VV&Ts, such as Formulated Problem VV&T, System and Objective Definition VV&T, Communicative Model VV&T, Programmed Model VV&T, Experiment Design VV&T, and Presentation VV&T appeared in the steps of the life cycle he proposed (figure 2.1). Also, during the formulation of a model, Data VV&T is required at each step, and Experimental VV&T is performed specifically between the experimental model and system and objectives definition.

Zenina and Merkuryev [2019], on the other hand, included model verification as a process that occurs during model validation. In their study, model validation is defined as the process to check how well the simulation model fits the observed system through comparison between of the logic and structure of the simulated system (in simulation world) and observed system (in the real world). In addition, model validation also covers the adjustment of calibration parameters and the assessment of model effectiveness.

The studies from Balci [1998] and Zenina and Merkuryev [2019] gave two different aspects of model verification and validation. From the aspect of Balci [1998], model verification and validation are two techniques to be performed in each of the stages through simulation model development. For Zenina and Merkuryev [2019], model verification is a part of model validation, where model validation means the whole justification of a simulation model.

By far, even the definition of model verification and validation can be concluded from the previous literature, the positions for both justifications in the development of a simulation model still remain unknown. To understand when and where to perform these two justifications, we look back to the research of Sargent [2010] again.

In his study, the meaning of model verification and validation was used from the terminology defined by Schlesinger [1979], where model verification is to ”ensure that the computer program of the computerised model and its implementation are correct” and model validation means ”substantiation that a computerised model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model”. The definitions are similar to the ones concluded in the previous paragraph. Something to be mentioned in this research is that there seems to be a standard for the naming of the justification process in the development of traffic simulation model. (figure 2.2).

For verification, there are specification verification and implementation verification. For validation, there are theory validation, conceptual model validation and operational validation. Each process mentioned above is to justify the relation between two objects during the development of the simulation model to ensure the quality and performance of the model. Therefore, even with different definitions, model verification and validation share a similar process of justification by ”comparing two objects”. The difference is that the word ”validation” is used in the process to justify a relation between two objects that one or both of them belongs to the real world, while the word ”verification” is a process to justify a relation between two objects that both of them are within the simulation world. To be more specific, the word ”validation” is used when there exists field data as a representation of ground truth, no matter the field data is used as an input to a model/theory or the observed phenomena itself. In model verification, only system data generated with intention by the developer or randomly generated to test the model is used.

In conclusion, even the definitions of model verification and model validation are not defined identically in different studies, each of them refers to similar meaning in most of the research. Both verification and validation are a process of justification that compare two objects in the development of a simulation model. The main difference is the use of the data, where model verification uses system data, while model validation uses filed data. This also decides when and where to conduct model verification or model validation. In this thesis, the definitions for model verification and validation to be used are as follows.

Model verification is a process to justify the implementation of the simulation model by comparing the simulation model with the conceptual model using system data. *Model validation* is a process to justify the performance, such as to what extent does the model represent the (real-world) system to be modelled, or the model in the simulation world by comparing the conceptual/simulation model with the indicated system using field data. The idea of conceptual model validation and operational validation from Sargent [2010] is therefore used in this study. With the result from the last subsection, a complete overview of the development of a traffic simulation model is presented in figure 2.4.

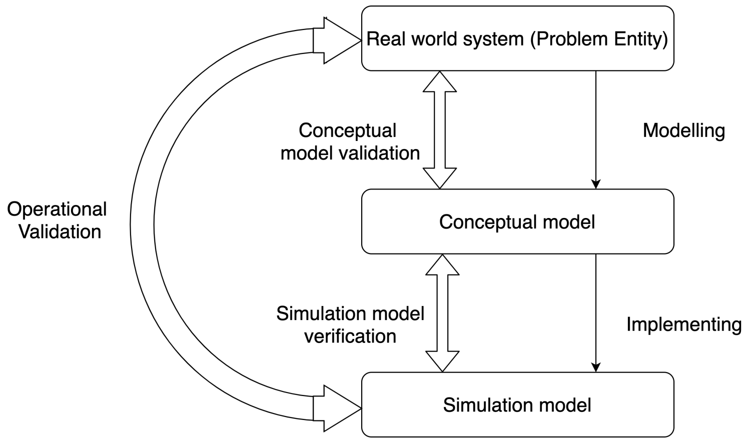

Figure 2.4: Brief schematic of simulation model development

In this subsection, the terminology of model verification and validation are well defined. With the definition, a clear overview of the development of a simulation model is illustrated. The focus of this thesis is on the model verification of traffic simulation model. It is now clear what model verification is. When and where to perform model verification are also indicated. In the next subsection, detail on how model verification is performed is exploited.

2.2 MODEL VERIFICATION

As the definition of ”model verification” is determined, the next step is to understand how to conduct a model verification. Similar to the definition of model verification and validation, there is no standard answer for model verification process.

It is not only because of the difference in the definition of model verification but also because the purpose to develop a simulation model could vary from model to model. The requirement for verification, along with the verification process is therefore different. In this subsection, the model verification process mentioned in previous studies is discussed. The objective is to find the critical steps and principles in common that are able to represent those methodologies used to verify traffic simulation model. As model verification is seldom discussed independently, literature on model validation procedures related to model verification defined in the last subsection are also included.

2.2.1 Model verification process

In the study of Taylor [1979], he performed model verification by using the four-level evaluation procedure introduced by Pilgrim [1975]. The four levels are: 1) Examination of the structure and the components of a model; 2) Estimation of model parameters; 3) Using another data set to exam the accuracy of the fitted model; and 4) Prediction of the applicable range of the model.

Benekohal [1986] explained that verification is performed through a series of sensitivity analysis. He mentioned that knowing the applicable range of the parameters of the model is important, so that the user/developer could know in which situation (not) to apply the model.

In the research of Kleijnen [1995], he conducted a survey on statistical techniques for model verification and validation. What he mentioned about the procedure is that the analysts should divide the model into modules and verify them one by one. There are four points in verification concluded from their study, which were 1) good programming skill, 2) check intermediate outputs through tracing and statistical testing per module, 3) examine the final simulation outputs by statistical tests against analytical results, and 4) use animation to inspect the dynamic behaviour of the simulation model.

Rakha et al. [1996] established a framework for systematic justification of traffic simulation models. In the study, they mentioned two objectives of model verification. First, ensure that, for a given input, the simulation model provides an output that is consistent with the conceptual model. Second, test the sensitivity of the parameters to confirm that outputs are consistent over a range of reasonable input. After that, they introduced a five-step schematic to perform tests in verification.

Balci [1998] explained that model verification is to build the model right. There are two points to check in model verification. First, to what extent is the conceptual model specified in the simulation model. Second, how accurate is the implementation of the conceptual model in the simulation environment. Instead of procedures, they provided 15 principles to follow when conducting verification, validation and testing (VV&T). One thing to be noticed is that in one of the principles, he mentioned that simulation model VV&T must be planned and documented, which can be observed in other researches, but only explicitly explained here. In his study, a model was also deemed as a composition of sub-models, and therefore, a sub-model was a unit to be tested and assessed.

Sub-model [Balci, 1998] and module [Kleijnen, 1995] are terms that often appeared in the studies related to model verification. The concept of module (sub-model), as a part of model, might come from the study of Parnas [1972], where the concept of module programming was first introduced. In the study, Parnas [1972] defined a module as a device with a set of switch inputs and readout indicators. He introduced a technique to specify (a) segments in a program precisely. There are four items to be specified according to his study, which are: 1) the set of possible values, which means the return values (outputs) of the module; 2) initial values of a function; 3) parameters; and 4) effect, as the mechanism of the function. This module specification is used for the implementation of a program. In addition to the items to be specified, Parnas [1972] also distinguished two classes of functions in a model. The first one is the function that provides information which cannot be determined without calling that function. The second one is named as *mapping function*, in which the value of the function is predictable from the current values.

In the research of Horiguchi and Kuwahara [2002], they concluded two types of vehicular traffic simulation model: macroscopic simulation model and microscopic simulation model. The features to be considered for the two types of model during model verification were summarised. In the study, model verification is to examine whether the simulation model can reproduce the result of the conceptual model, on which the simulation model is based. They suggest using system data to perform verification of simulation model instead of field data so that each function of the model can be tested separately. To be more specific, they explained the process of model verification for different types of model. For macroscopic simulation models, the verification process is like a self-consistency check to assess whether the simulation result agrees with the given parameters. For microscopic simulation models, model verification is to recognise the relationship between model parameters, concerning individual road user, and model behaviour reproduced by the simulation. Therefore, it is more like a sensitivity analysis to understand the influence of major parameters on model behaviour.

Lastly, in the study of Sargent [2010], he discussed on both model verification and validation process, but in different ways. For model verification, it is considered to determine the correctness of the implementation of the simulation model. In the study, a simulation model was said to consist of two components, simulation functions and computerised conceptual model. Simulation functions are the elements especially used to run a simulation model such as time-flow mechanism, random number generator, and random variate generators. The computerised conceptual model is the mathematical or logical representation of the problem entity that being implemented in the simulation environment using a programming language. The model verification is, therefore, to ensure that the conceptual model is implemented correctly that it can be run in a simulation environment. The procedure described in the study is that each sub-model is tested first to see if it works properly, and the whole model is executed with different scenarios to explore the relation of the inputs and outputs.

For model (operational) validation discussed in the study of Sargent [2010], it is to determine whether the performance of the simulation model has the accuracy in the study area that the model is proposed to. Although this complete procedure is used for operational validation, it is still worth studying since the only difference between model verification and model validation defined in the previous subsection is the objects to be compared when performing justification. By switching the problem entity used in the research of Sargent [2010] with the conceptual model defined in this research, the eight steps can also be deemed as a process of model verification. The eight steps in model validation explained in his study are:

1. The user/developer specifies validation approach and decides the validation techniques to be used in the process.

2. The user/developer specify the level of accuracy required of the performance of the simulation model according to the intended application of the model.

3. Test the assumptions and the theories the simulation model based on whenever there is a chance to do so.

4. Perform face validation on the simulation model in each run of the simulation.

5. Explore the model behaviour of the simulation model in each run of the simulation.

6. Make comparisons between the simulation model and problem entity using output data in a different scenarios.

7. Record the validation and make documentation.

8. Make a schedule for periodic review of model validation if a long term use is expected for the simulation model.

From the past research, a few points related to model verification process can be concluded. First, there are two practical objectives in the model verification process. One is to ensure the basic functionality of the simulation model by reproducing the result of the conceptual model which the simulation model is based on or by performing intended application of the simulation model [Horiguchi and Kuwahara, 2002; Rakha et al., 1996; Sargent, 2010]. The other one is to understand the input-output relation in the simulation model in order to determine the applicable range of the simulation parameters. This process often includes a sensitivity analysis [Taylor, 1979; Benekohal, 1986; Horiguchi and Kuwahara, 2002; Rakha et al., 1996; Sargent, 2010]. These two objectives are often discussed together.

However, from the author's perspective, these two objectives belong to two different levels of model verification. In level 1, as the first objective, the verification is performed to ensure the basic functionality of the simulation model. Therefore, the verification process focus on the comparison between the conceptual model and the simulation model. In level 2, as the second objective, the verification is to decide the applicable range of the parameters. Because of that, the sensitivity test within the simulation model is the main action, with the support of comparing to the conceptual model. The sensitivity tests cannot be carried out before the basic functions are confirmed. Therefore, the verification of level 1 must be conducted before level 2. The model passed the verification of level 1 could be said verified in the study area where the conceptual model is proposed, while the model passed the verification of level 2 could be deemed as verified in the study area within the recognised range of the parameters.

Second, verification is often performed per module or sub-model [Kleijnen, 1995; Balci, 1998; Sargent, 2010], and this enables a justification in an organised way. As a simulation model is usually composed of several functions, testing the whole model directly could be dangerous with the correctness of each module unknown. Also, if a simulation model fails in a certain test of verification, it will be easier to find the problem when the user/developer already knows which part of the model is being examined. Because of that, to perform model verification, the modules inside the simulation model need to be identified beforehand.

According to Parnas [1972], the basic elements to be specified in a module are the effects, the parameters, and the outputs of the module. Besides, the relationship between the modules, such as an output of a module being a parameter used in another module, and the type of a module should also be indicated from the author's perspective since the relationship would decide the order of testing and the type of a module would affect the techniques used.

This action, identification of module, was not notably described in the past studies, since it is similar to the model specification during the implementation of the conceptual model [Sargent, 2010]. However, the specification is known only when the verification is performed by the developers who implemented the simulation model themselves. If the verification process is performed by a third party, which can reduce the bias during the development of a simulation model [Balci, 1998], the structure and the components of the simulation model need to be analysed first to carry out an examination.

Third, modules are categorised into three types: simulation function, computerised conceptual model [Sargent, 2010], and mapping function [Parnas, 1972], where simulation function is related to the execution of a simulation, computerised conceptual model is the programmed representation of the model, and mapping function is to make the code clear. To avoid the confusion of similar terms used in this research, ″simulation module″ will be used instead of ″simulation function″, and ″mapping module″ will be used instead of ″mapping function″ in this research.

Last, a plan is a necessity to perform model verification. Nevertheless, the importance to make a plan before performing tests of verification was only discussed by Balci [1998]. The plan should consider the structure and the components (modules) of the model, the aimed level of verification, along with the techniques to be applied to achieve the purpose of model verification.

In conclusion, model verification is not only about testing, but also about obtaining the information of a simulation model and planning for the testing. Therefore, to create the framework for model verification, it should include the two phases in verification, ″Creation of Plan″, where the information of the model is collected to make the plan, and ″Execution of Plan″, where the tests according to the plan are performed. In phase ″Creation of Plan″, there are three steps to be completed, which are 1) determine the level of verification; 2) identify the modules that build up the simulation model; and 3) make a plan based on the information obtained. On the other hand, in phase ″Execution of Plan″, the action is about performing the tests indicated in the plan in an organised way. Model verification is often deemed as a series of tests to ensure the quality of the model, and the preparation for the tests is easily overlooked. However, these preparation are critical steps in model verification, guaranteeing the tests performed in the right direction to minimise the bias and errors that emerged during the development of the simulation model.

In this subsection, two phases along with their critical steps in model verification are summarised and constructed from the literature. As a plan is necessary, the content of a plan is to be found out. In the next subsection, the discussion is made on what to be executed in a plan, basically the approaches and the techniques used to verify a traffic simulation model.

2.2.2 Model verification approach

In this subsection, the approaches and techniques applied to verify traffic simulation model are discussed. Model verification and validation share similar approaches since both of them consider justification as a process through comparison. Because of that, the approach of model validation from past studies that related to model verification defined in this study are also included.

In the research of Benekohal [1986], he verified and validated a car-following model he developed. The verification process was a series of sensitivity analysis to find out the applicable range of the model. On the other hand, the validation he practised were graphical comparison on speed profile and trajectories (from simulation results and field data). He analysed the difference in the graph and drew the conclusion of whether the model was validated.

In the study of Kleijnen [1995], model verification is done by three actions. For the intermediate outputs come from modules, they suggested tracing and face validation. To compare final simulation outputs with analytical results, statistical techniques are the primary choice. In case the model is a dynamic system, animation is advised to investigate the dynamic behaviour of the model. However, they mentioned that animation could be dangerous since the developer/user tend to focus on short simulation runs, so the problems that occur in long runs is overlooked. For model validation, statistical analysis and regression analysis on the simulation outputs and field data was recommended. Sensitivity analysis was said to apply when exploring the model behaviour with inputs and outputs that cannot be observed. Also, risk analysis, such as Monte Carlo sampling, could be used to test the robustness of the model.

In another study, Horiguchi and Kuwahara [2002] summarised six basic features to verify a vehicular traffic simulation model. The features are related to the findings and critical values based on the fundamental theories of vehicular traffic flow behavior. For both macroscopic and microscopic simulation model, these six features are to be checked in model verification. The difference in approach for macroscopic and microscopic simulation model came from the fact that macroscopic simulation model explicitly, or implicitly, implies the capacity to a section of the road, while microscopic simulation model does not have a parameter indicating the capacity of a link. Because of that, the approach for verifying macroscopic simulation model was consistency checks through the six features while they use sensitivity tests for verifying microscopic simulation model.

Balci [1998] introduced a vast number of techniques to be used in verification, validation and testing (VV&T). Since VV&T in this study is a bigger concept that indicates the justification in each step of model development, the techniques they mentioned were not specific for model verification or model validation defined in this research. Some of the techniques were related to the process of communication between different groups participated in the development of the simulation model, while some were used especially for programming skill. They identified the techniques into four categories: informal, static, dynamic, and formal. The techniques that are relevant to model verification and validation defined in this study are face validation (where people who know the system subjectively compare model and system and judge whether the model outputs are reasonable), functional testing (which is to evaluate the accuracy of model input-output transformation), graphical comparisons, sensitivity analysis, special input testing, statistical techniques, and visualisation/animation.

Lastly, Sargent [2010] also mentioned a number of techniques to be used for model verification and validation. For model verification, the two techniques that can be used for investigation of each module and the whole model are traces and face validation. In traces, the behaviour of different types of entities is traced through each module and the whole model to determine if the model's logic is correct. Face validation was defined similarly to the previous literature. For model validation, they mentioned similar techniques that appeared in other literature, such as animation, comparison to other models (comparing results from other "valid" models), face validation, graphical comparison, sensitivity analysis, extreme condition tests, and traces.

One thing to be noticed is that, besides the techniques, Sargent [2010] introduced a measure to decide which type of techniques to use based on decision approach and system observable (table 2.3). Decision approach could be subjective approach or objective approach while system observable was determined whether the data of the system can be collected.

Table 2.3: Classification of approach proposed by Sargent [2010]

	Observable System	Non-observable System
Subjective Approach	• Comparison Using Graphical Displays • Explore Model Behavior	• Explore Model Behavior • Comparison to Other Models
Objective Approach	• Comparison Using Statistical Tests and Procedures	• Comparison to Other Models Using Statistical Tests

Sargent [2010] classified the aforementioned techniques into comparisons of output behaviours (including comparison using graphical displays and comparison using statistical tests and procedures), explore model behaviour (which usually implies sensitivity analysis and face validation), comparison to other models and comparison to other models using statistical tests. He mentioned that for an observable system, a comparison using statistical tests is always preferred since it allows for objective decisions. However, a graphical comparison is sometimes needed when it is not possible to use statistical test, such as the condition that statistical assumptions being hard to be satisfied.

From the previous research, there are a few things to be noticed. First, the techniques applied in the past studies are quite in common. They are summarised with definitions as follows:

Animation: Displaying graphical images of the internal and external dynamic behaviour of a simulation model during (or at the end of) execution which enables the user/developer to discover errors observing with related knowledge. [Balci, 1998]

Consistency check: When the result is predictable, as the capacity of the macroscopic simulation mentioned by Horiguchi and Kuwahara [2002], which is an input to the simulation model, the user/developer used the input as a reference value to examine whether the simulation model generates output(s) correspond to the input.

Face validation: Experts in the subjects or the system are asked to evaluate whether the model and its behaviour are reasonable [Sargent, 2010].

Functional testing: It is used to assess the accuracy of model input-output transformation by feeding inputs (test data) to the model and evaluating the corresponding outputs [Balci, 1998]. The process is similar to the risk analysis mentioned by Kleijnen [1995], to evaluate the robustness of the model, which could also be done by Monte Carlo sampling.

Graphical comparison: The graphs of values of model variables over time are compared with the graphs of values of system variables to investigate characteristics such as similarities in periodicity, skew, number and location of inflexion points, logarithmic rise and linearity, phase shift, trend lines, and exponential growth constants [Balci, 1998].

Sensitivity analysis: This technique consists of changing the values of the input and internal parameters of a model to determine the effect upon the model's behaviour or output [Balci, 1998].

Special input testing: It is used to assess model accuracy by testing a variety of inputs to the model Balci [1998]. The process might look similar to sensitivity analysis. However, the objective is not finding out how sensitive are the parameters; instead, it is used to test the robustness of the model by giving extreme inputs or invalid inputs.

Statistical tests: The use of confidence intervals or the use of hypothesis tests to make an objective decision, including regression, comparison, correlation or other non-parametric tests [Sargent, 2010].

Tracing: The use of tracing is the tracking of model behaviour through each module and the overall model to determine if the logic is correct and if the necessary accuracy is maintained [Sargent, 2010].

Second, the verification process and approach depends on the target indicators. Benekohal [1986], Horiguchi and Kuwahara [2002] explicitly described the target indicators to carry out the comparison, which makes it clear where to use the techniques. Both of the studies indicated the importance to understand the type of the model to be verified. They shared similar indicators since they both discussed vehicular traffic simulation model.

Third, the classification of approach from Sargent [2010] provides an unambiguous measure to decide which types of techniques to apply when conducting verification and validation. As two levels of model verification have been recognised in the last subsection, the techniques from the past studies can also be categorised into these two levels. By combining the classification from Sargent [2010] with the level of verification we concluded in the studies, it is possible to separate the techniques to apply when verifying a model in different level of verification. As a result, the conditions when and where to apply a certain technique for what reasons become understandable as shown in table 2.4.

Table 2.4: Techniques to be selected for the level 1 and 2 of model verification. (The highlighted techniques with bold text should be the main focus at the level.)

	Subjective Approach			**Objective Approach**
	Compare Graphical Display	*Explore Model Behavior*	*Other*	
Lv.1	Graphical Comparison	Animation Tracing	Consistency Check Face validation	**Statistical tests**

	Subjective Approach			**Objective Approach**
	Compare Graphical Display	*Explore Model Behavior*	*Other*	
Lv.2	Graphical Comparison	Animation Functional testing **Sensitivity analysis** Special input testing Tracing	Face validation	Statistical tests

In this subsection, model verification approaches and techniques are discussed. The techniques used to verify traffic simulation model are summarised with explanations. The decision process to select a certain approach is also elaborated. By combining the measure to decide techniques to use with the classification of different purpose of model verification concluded in the previous subsection, a complete table describing when and where to use a certain verification technique for what reasons is built.

In the next subsection, the conclusion to the first part of the literature review is made. The results obtained at each subsection are inspected and summarised.

2.2.3 Conclusion to model verification

In these two sections, a comprehensive discussion on traffic simulation model verification is made. First, the three stages in the development of traffic simulation model are identified, which are Problem Entity, Conceptual Model, and Simulation Model, along with the two processes from one stage to another stage, which are modelling and implementation. The justification processes, model verification and model validation are then indicated. To clarify the difference between two justification processes, the terminology of model verification and validation are discussed and defined. With a clear definition, an overview of the development of a simulation model is illustrated.

To give an insight into model verification, literature regarding the process of model verification is examined. Two practical objectives are summarised from previous studies, representing the two levels of model verification. Two phases with according steps are identified in the model verification process, which are phase "Creation of Plan" with steps 1) determine the level of verification, 2) identify the modules that build up the simulation model, and 3) make a plan to perform model verification and phase "Execution of Plan", containing the action to perform the tests in a plan. Lastly, to provide more detail on how to execute the plan, model verification approach and techniques are discussed. The techniques that are used to verify simulation model are summarised with the decision process on how and when to select a certain approach.

The result of the first part of the literature study is a concept of the framework derived from the methodologies used previously with clear definition and process to follow. The framework includes two phases and the according steps, along with the explanation of the detail in each step. The framework is to be thoroughly discussed in the following chapter.

In this part, the relationship between model verification and cyclist simulation model is not yet discussed. From Horiguchi and Kuwahara [2002] and Benekohal [1986], we found out that no matter for which level of model verification, the indicators used to perform verification is important. Bike traffic shares some traffic flow variables with vehicular traffic. However, with a higher degree of freedom in movement, cyclists behave differently from vehicles. The definitions of the traffic flow variables are sometimes different even when using the same terminology, and there might need more variables to describe the behaviour of a cyclist precisely. Because of that, the indicators used to perform cyclist model verification might be different and remained to be discovered.

In the next part, the literature related to the microscopic cyclist simulation models are to be discussed. The focus is on microscopic simulation model, in which the difference between cyclist simulation model and vehicular simulation model, such as steering and accelerating behavior, are presented explicitly in the model. The objective is to look into the indicators used to describe the model behaviour in order to explore the potential difference in verification of cyclist simulation model. Based on that, the connection between model verification and cyclist simulation model can be made.

2.3 MICROSCOPIC CYCLIST SIMULATION MODELS

The microscopic modelling of cyclist has the following characteristics. The cyclist is modelled individually, and each of them has its own properties. Each cyclist has its objective, and each of them responds to its surroundings, which includes physical environment, other road users, and information environment. Different from vehicular traffic, the behaviour of bike traffic is less anisotropic, which means bike traffic has a higher degree of freedom in its movement and this characteristic is explicitly described in the microscopic simulation models. In this section, the microscopic cyclist simulation models are introduced, including the cyclist queue formation model mentioned in the introduction. The objective is to understand the parameters used and the according outputs of the models and they are summarised in table 2.5.

Taylor and Mahmassani [1998] used a discrete choice model to estimate the factors that impact the gap acceptance behaviour of cyclists and motorists in mixed-traffic. The results indicated that both cyclists and motorists accept smaller gaps when the next coming road user is a cyclist, and both require larger gaps if a larger vehicle is the next coming road user. The parameters used in the model is the type of the decision-maker (cyclist, motorcyclist), gap length (in time), and the vehicle type on the closing gap (cyclists, motorist, vehicle, bus, truck). The intermediate output of the model is the probability that a certain gap is accepted by giving the aforementioned attributes. The output that accumulating the intermediate outputs with parameters from field data is the probability distribution of the accepted gap, and the centre of this distribution is often chosen as the critical design gap to conduct traffic management.

Vasic and Ruskin [2012] used a rule-based model, cellular automata, to simulate mixed traffic, including cars and bicycles for two scenarios. In cellular automata, space and time are discretised, and cyclists have a set of rules to follow while moving through space. They addressed the mixed traffic of vehicle and bicycle with a technique to model network infrastructure and network occupancy (of vehicles) based on one-dimensional cellular automata components. The parameter used in the model is the layout of the intersection (for spatial modelling) and the inflow of both vehicle and cyclist. The output of the model is the capacity at the intersection with left-turn traffic flow.

Liang et al. [2012] proposed a psychological-physical force model to represent bicycle dynamics. A force model obeys to Newton's second law, so the acceleration of an object can be calculated with the mass of the object and force on the object given. By giving values to the parameters of those forces along with the current state of a cyclist, the next state, e.g. speed, acceleration and position, can be calculated. The intermediate output is the dynamic of a cyclist, representing as the direction and acceleration. By accumulating the result with simulation, a final output as a fundamental diagram indicating the speed-density relationship of the cyclists is obtained.

Jiang et al. [2013] studied the deceleration, acceleration and crossing behaviours of bicyclists at signalised intersections. They used statistical analysis and nonlinear regression methods to develop a velocity model for acceleration and deceleration behaviour of cyclists when departing and approaching the stop line. A discrete choice model was also used to estimate the gap acceptance behaviour of cyclists when crossing against right-turning vehicular traffic at signalised intersections. They found that cyclists started to decelerate when they are within 30 m from the stop line and that their acceptance of a gap depends on the speed of the cyclist and the speed of the vehicle, as well as the size of the available gap. The parameters used and the outputs of the gap acceptance model are similar to the one proposed by Taylor and Mahmassani [1998]. The parameter used in the velocity model is the current states of the cyclist, which includes whether it is departing or approaching to the stop line, the distance to the stop line and the present speed. The output is the speed profile and the acceleration profile of the cyclists against time and position.

Ma and Luo [2016] used naturalistic data to derive the cyclist acceleration process without interacting with other road users. The U-shape curves were first observed for both acceleration and deceleration processes of cyclists. The polynomial model was then used to form three acceleration models that represent the U-shape curved, and they were estimated using the maximum likelihood method. While the model with more parameters shows superior performance, the simplified ones were still capable of catching the trends in the acceleration profiles. Effects of social-economic characteristics were also investigated. They found out that gender and bike type (agility) had significant effects on the acceleration decision. The parameter used in the model are the ones estimated with the field data and the attributes of the cyclists (gender, agility), and the output of the model is the acceleration and the speed of the cyclist across the time.

Huang et al. [2017] proposed a social force simulation model for cyclist behaviour with heterogeneous traffic at unsignalized intersections. They modelled the force from other road users, such as vehicles, cyclists and pedestrians. They found out that avoiding collisions with other road users is one of the characteristics of cycling behaviour at intersections. The simulation model is verified with the trajectories and the speeds from observation and simulation. Both trajectories and speeds from simulation exhibited no significant difference to the empirical data, which means that the bicycle crossing model can reasonably reflect the cyclist crossing behaviour characteristics at unsignalized intersections. Similar to Liang et al. [2012], they used a force model to represent the interaction of road users as a combination of force, which results in the acceleration of an object. Therefore, the parameters used in the model is the attributes of the road users at the intersection. The intermediate output is the resultant force (which become acceleration) to a cyclist according to the existence of other road users at the moment. The aggregate result across a time period is speed profile and trajectories of cyclists.

Gavriilidou et al. proposed a novel cyclist operational model to describe the queue formation process of the cyclist at a signalised intersection. A two-layer framework of discrete choice model is used to describe the operational decision-making process at different levels, mental and physical. In the mental layer, cyclist's preference for the stop position at an intersection is described. The important factors that influenced the stop position of the cyclists, such as the distance to the stop line, the distance to the prior cyclist, the existence of the request-green button, were estimated. The parameters used in the model include the attributes of the cyclist and the attributes of each stop position (which is discretised into diamond cells to represent the stop space of a cyclist). The intermediate output is the probability of each stop position that the coming cyclist would stop. The aggregated output across the time period with a number of cyclists is the probability distribution of the cyclist stop (queuing) positions.

Given the stop position, in the physical layer, the preference to change in steering angle and speed of a cyclist is illustrated. The factors that affect the cyclist to steer and to accelerate/decelerate, such as the existence of the prior cyclist, the distance to the stop line, and the distance to the aimed position, were estimated. The choice set of the physical layer discrete choice model was represented as a circular sector. The alternatives were discretised into cells, and each cell described a combination of steering angle and speed change. The parameter used in this model were the current state and the predicted state (if a certain cell is chosen) of the cyclist (current steering angle, speed, position etc.), and the interacting environment, such as the state of other cyclists. The intermediate output was the probability of a certain combination of steering angle and speed change that a cyclist would perform. The aggregated output across the time period with a number of cyclists is the probability distributions of the combinations of steering angle and speed change that cyclists would perform.

To conclude, many types of model have been applied to describe the behaviour of the cyclists, including the mathematical polynomial model [Ma and Luo, 2016] that directly depicts the speed curve, or models that used in other fields, such as rule-based model [Vasic and Ruskin, 2012], velocity-based model [Jiang et al., 2013], force-based model [Liang et al., 2012; Huang et al., 2017] and utility-based model [Taylor and Mahmassani, 1998; Jiang et al., 2013; Gavriilidou et al., 2019b]. The parameters used in the cyclist simulation model are usually the current state, which describes the dynamics of the cyclist, or the attributes, including some general information of the cyclist.

The outputs of a simulation model can be categorised into two types, the intermediate outputs and final outputs. The intermediate outputs come from the logic of the conceptual model. For example, the probability to an alternative from a discrete choice model is an intermediate output; the resultant force to an object from a force model is also an intermediate output. These intermediate outputs are used to describe a decision or a dynamic of a cyclist at a certain point. However, one intermediate output at a certain moment doesn't mean much. The final outputs, accumulated from the intermediate outputs, are the product that represent the behaviour of cyclists. In the past studies, speed profiles and positions (trajectories) are often used as the final outputs. Sometimes a probability distribution is also used to represent a specific behaviour of cyclists, such as gap acceptance and queuing behaviour. Liang et al. [2012] are the ones that tried to construct a fundamental diagram indicating the speed-density relationship of the cyclists.

To this point, it is still hard to conclude the potential difference to verify a cyclist simulation model. Therefore, we take a look to the microscopic vehicular simulation model. In one of the researches, Hoogendoorn and Bovy [2001] summarised the vehicular traffic simulation models in the past fifty years. According to the research, the majority of the studies on microscopic vehicular simulation models were car-following models, since the research efforts focused on that during 1960s and lasted for decades. Car-following models are based on the assumption that a vehicle will maintain a minimum gap to the proceeding one, describing the process of one vehicle following another. They explained that car-following models have been mainly applied to single lane traffic since the lane-changing processes cannot be easily described. Because of that, the capacity of the stream can be analysed using car-following models. In terms of describing the lane-changing behaviour, the car-following models usually incorporate other models, such as gap acceptance models, to separate the decision to perform a lane-change.

In the previous studies of cyclist models, bike traffic are also categorised into different parts. The gap acceptance behavior is also one of them. However, the gap acceptance models for cyclists are usually used to describe the crossing or turning behaviour [Taylor and Mahmassani, 1998; Jiang et al., 2013] instead of lane-changing behavior. To be more specific, it is unusual to distinguish the horizontal behavior from the longitudinal behavior for cyclists, since there is not a strict lane for cyclists to line up one after one. Because of that, bicycle dynamics in both directions (horizontal and longitudinal) are often modeled together [Vasic and Ruskin, 2012; Liang et al., 2012; Huang et al., 2017; Gavriilidou et al.].

Based on this, we concluded that the possible difference to verify the two types of model could come from the dimension of the modeling behavior. Due to the anisotropic of vehicular behavior, it is more intuitive to justify whether an output is reasonable for vehicular model than cyclist model. For example, the headway (in distance) is usually considered in longitudinal direction only for vehicular models, but both directions need to be taken into account for cyclist models.

Also, according to Hoogendoorn and Bovy [2001], there were 58 microscopic vehicular simulation models identified by Algers et al. [1997] at the time. With the majority of them discussing the car-following behaviour, the large number of models under similar assumptions make the comparison between models more possible to be carried out. In terms of the justification of a model, this could be an advantage for vehicular models when a verification of a simulation model is to be conducted with its conceptual model inaccessible.

In this section, the microscopic cyclist simulation models are discussed, including the methodologies applied, the parameters used and the according outputs of the model. The difference between the cyclist simulation model and vehicular simulation model is briefly addressed. From the author's perspective, there is not really an influence of this difference on the strategy to perform model verification at a cyclist simulation model. When it comes to a justification by comparing different models, the use other model(s) as a reference to the studying behavior could be limited for cyclist models. However, it is also the same for the newly developed vehicular simulation model if the assumptions are made differently to the car-following behavior. In the next section, the conclusion to the literature review is made. The connection between model verification and cyclist simulation model is explicitly discussed.

2.4 CONCLUSION TO LITERATURE REVIEW

In the first two sections of the literature review, a comprehensive discussion on simulation model verification is made. The definition of model verification is defined as "a process to justify the implementation of the simulation model by comparing the simulation model with the conceptual model using system data". A brief overview of model development is built with the three stages (problem entity, conceptual model, and simulation model), two model developing processes (modelling and implementation) and two justification processes (verification and validation). Model verification, as one of the justification process, is thoroughly discussed on its procedure and approach. The two phase and according steps of model verification are concluded along with the measure to determine the approaches to perform verification.

In the third section, an overview of the microscopic cyclist simulation models is deliberated. The parameters used and the according outputs of each model are examined. There are usually two types of outputs from a simulation model. One is the intermediate output, which often comes from the logic of the conceptual model, such as the probability to an alternative or the resultant force to an object. These intermediate outputs describe the movement of the cyclist at a certain point, but they do not mean much individually. The accumulation of the intermediate outputs, as the final output(s), is the one that represents the behaviour of cyclists. From the second part of the literature review, it is concluded that speed profiles and trajectories are often used as the final outputs of a microscopic cyclist simulation model, representing the individual behaviour of cyclists.

Aside from the parameters and outputs of the simulation model, the difference between cyclist simulation model and vehicular simulation model is also mentioned.

The relative small dimension of output to describe the modeling behavior for vehicular model enables an intuitive judgement of the outputs. Also, the large number of vehicular studies under similar assumptions makes it possible to explain a result by comparing to other similar models. Therefore, a "second opinion" from the comparison to another model could be helpful to assess whether the output is reasonable or not. Although this difference could limit the approach of comparing different models to verify a cyclist simulation model, it does not have much influence on the whole verification methodologies we concluded in this chapter.

At the end of this section, the suitability of the content of the literature review is discussed, including a conclusion that answers the sub research questions of this chapter. The first sub-question of this chapter is:

Q1-1 *What is model verification for traffic simulation model?*

The answer to this question can be separated in two parts:

Q1-1.1 What is the definition of model verification?

Ans: Model verification is a process to justify the implementation of the simulation model by comparing the simulation model with the conceptual model using system data.

Q1-1.2 What is the purpose to perform model verification?

Ans: There are two objectives of model verification: 1) ensure the basic functionality of the simulation model; and 2) decide the applicable range of the simulation model. These two objectives result in the two levels of model verification where level 1 of verification focuses on the comparison between conceptual model and simulation model while level 2 focuses on the sensitivity test within the simulation model.

The second sub-question of this chapter is:

Q1-2 *What methodologies that were used to verify vehicular traffic simulation model are found in literature?*

The three supporting questions are used to answer this question:

Q1-2.1 What was the process to perform model verification?

Ans: There are two type of verification process found in the literature. The first type is a series of tests assessing the functionality of the simulation model while comparing the simulation model with the conceptual model (of the simulation model). This type of process usually starts with testing the modules (a part of the model) first, then assessing the whole model after that. The second type is a series of sensitivity tests on the crucial parameters assessing influence of the parameters on the model performance.

Q1-2.2 What were the approaches applied in model verification?

Ans: The approach found related to the model verification in the previous studies are: animation, consistency check, face validation, functional testing, graphical comparison, sensitivity analysis, special input testing statistical test, and tracing. Among the approaches, statistical test is found the most promising approach to complete level 1 of verification with objective decisions and sensitivity analysis is deemed as a necessity to perform level 2 of verification.

Q1-2.3 How were the process and approaches of model verification determined?

Ans: The process is basically decided by the purpose of verification. For level 1 of verification, the process is often aligned with the first type of process mentioned above. Similar for level 2 of verification, the process usually follows the second type of process. On the other hand, the approaches are found determined by three factors. One is the purpose of the verification, one is decision approach, and one is system observable.

The third sub-question followed by its answer is then presented:

Q1-3 *What are the potential differences to verify a cyclist simulation model from verifying vehicular simulation model?*

Ans: Based on the findings in literature review, there is no decisive difference between verifying cyclist simulation model and verifying vehicular simulation.

The last sub-question in this part along with its answer are:

Q1-4 *What are the steps to be included in the structure of the methodology to represent a verification framework for traffic simulation model based on the findings from the previous methodologies?*

Ans: In the literature review, it is found that there are actually two phase to perform model verification, one related to the preparation while the other focus on the execution. Therefore, the two phase "Creation of Plan" and "Execution of Plan" is constructed first. In phase "Creation of Plan", there are three actions to be completed according to the previous studies. The first action is to understand the purpose of the model to be verified so that the level of verification can be determined. The second is to understand the structure of the simulation model, in which the modules in the model and the whole model to be verified are recognised. The third is to construct a plan based on the information retrieved from the first two actions. By putting these actions into an appropriate order, we can conclude the three main steps to be completed in the first phase: 1) "Determine level of verification", 2) "Module Identification" and 3) "Verification Planning".

For phase "Execution of Plan", the main action is to perform the tests indicated in the plan. However, when it comes to a failure in a test, there should be actions to find and fixed the problem(s). Therefore, the first action to be included besides the main action is analysing the problem. From the understanding of the development of a simulation model, the problem could happen in modeling (from system to conceptual model) or in implementation (from conceptual model to simulation model). Also, there could be mistake when making the plan. Depending on where the problem belongs to, the following actions are included. First, if the problem is in the plan, then a modification of the plan is required. If the problem belongs to the implementation, a modification of the simulation model is needed along with a modification in plan due to the possible change made in the simulation model. Last, if the problems belongs to modeling, probably a modification of conceptual model should be performed. However, as model verification is defined to justify the implementation process, it is assumed that the problems in modeling should have been fixed before implementation. Therefore the action to fix the problem in conceptual model is not included. To summarise, the main steps to be performed in the second phase are: 1) Perform tests, 2) Analyse problem (when failing in a test), 3-1) Modify simulation model, and 3-2) Modify Plan.

This part of this study is finished as the sub-questions are answered. In the next chapter, a thorough discussion is made on the framework to be applied to verify a cyclist simulation model constructed based on the information collected in this chapter. The two phases and the corresponding steps, including the beginning procedure, is exploited. The additional information related to the verification approach and cyclist simulation model characteristic will be the support to explain the framework.

Table 2.5: Summary of the cyclist simulation models

	Method	Target behaviour	Indicator used	Intermediate result	Final result
Taylor and Mahmassani [1998]	Discrete choice model	Gap acceptance behaviour	Type of the decision-maker, gap length, vehicle type.	Probability to accept a gap	The probability distribution of the accepted gaps
Vasic and Ruskin [2012]	Rule-based model: cellular automata	Bicycle dynamics	Layout of the intersection and the inflow of vehicle and cyclist		The capacity at the intersection
Liang et al. [2012]	Psychological-physical force model	Bicycle dynamics	Forces attributes, speed, acceleration, position	Resultant force	A fundamental diagram indicating the speed-density relationship
Jiang et al. [2013]	Nonlinear regression Discrete choice model	Bicycle dynamics Crossing behaviours	States of the cyclis, distance to the stop line and speed.		Speed profile and acceleration profile of the cyclists against time and position.
Ma and Luo [2016] used	Polynomial model	Bicycle dynamics: Cyclist acceleration process	Estimated parameter, cyclists attributes		Speed profile and acceleration profile over time.
Huang et al. [2017]	Social force model	Bicycle dynamics	Attributes of road users	Resultant force	Speed profile and trajectories
Gavriilidou et al. [2019]	Discrete choice model	Queue formation process	Attributes of cyclist, attributes environment, cyclist state (speed, position)	Probability of stop position Proobability of steering and pedalling	Probability distributions of the cyclist stop (queuing) positions Probability distributions of the cyclist steering angle and speed change over time

3 VERIFICATION FRAMEWORK

In this chapter, the development of the verification framework is discussed based on the two phases and the according steps summarised from the literature review. The two phases are "Creation of Plan" and "Execution of Plan". To do so, the primary purpose of the framework is explained first. A brief description of the framework is then presented. Following that, each step in the framework is elaborated with the reason why that particular step should be performed and how it should be performed when applying the framework to a traffic simulation model. In the last section of this chapter, a review of the framework will be addressed.

3.1 OVERVIEW OF THE FRAMEWORK

Before introducing the structure of the framework, the author would like to present the purpose of the framework, as well as the expected outputs from the framework. Therefore, in this section, why the framework should be created and how the framework look like will be discussed.

3.1.1 Purpose and product

As mentioned in the literature review, verification is an important step in simulation studies as a necessary process to develop a simulation model. The experience of verifying a model is valuable to be studied. However, from the past literature, we can see that the approaches to verify a traffic simulation model could be very different. It could be several examination tests of simulation models using virtual data to confirm their fundamental functions [Horiguchi and Kuwahara, 2002], or it could be structured walk-through and traces [Sargent, 2010]. No matter which approach is used to verify a simulation model, it is usually hard to adopt the past experience to verify another simulation model since the process, tests, and indicators used in the previous model might be different from the model to be verified. So, it is needed to create a framework as a standard procedure for model verification that users/developers can follow in a systematic way and their experience can be shared and studied with a comparable format, at the same time ensuring the quality of verification.

The final product from the framework is expected to be a verified simulation model and the completion of the verification can be ranked into two levels as concluded in literature review, level 1 and level 2. Besides the final product, the identification of which parts of the input model are functioning as expected and which parts are not is also an important output of the framework. By completing level 1 of verification, the basic function of the simulation model is assured by comparing with the conceptual model, and by completing level 2, the applicability of the simulation model across a reasonable range is confirmed through sensitivity analysis. These quantified definitions of the quality of a verified simulation model is introduced to assess the verified simulation model, and thus, the verified simulation models can be compared with each other.

Also, the procedures to verify a simulation model can be traced through the framework so that the other model users/developers can understand what is the decided level of verification, what are the elements inside the model, why certain techniques are used to verify the model, and how tests were performed in verification. Therefore, the framework enables the users/developers to look back how a simulation model is verified. To conclude, the products are a verified simulation model with its quality explicitly defined, an identification of required modification and a comparable and traceable verifying procedure.

3.1.2 Structure

After the introduction to the purpose and the product of the proposed framework, this part of the section is dedicated to the introduction of the structure of the framework. By visualising the framework, it is able to indicate the consequences of the framework. As a result, it is possible to determine how the framework should be applied. The framework is initially constructed based on the literature review of the verification process. The two phases and according steps summarised from the previous research are combined together as presented in figure 3.1.

The start of the framework is to have a simulation model as an input to see whether it fits this framework, so to help the user to know if he/she can use this framework to verify the model (section 3.2). Basically, the model is checked on whether it fits this framework with some basic questions. If it does, the framework may be applied to the case, and the user can continue to the framework. In phase 1, "Creation of Plan" (section 3.3), the first step is "Determine Level of Verification" (3.3.1). The action of this step is to understand the purpose of the simulation model so that the user can determine the aimed level of model verification of the input simulation model. Step 2, "Module Identification" (section 3.3.2) is to inspect the simulation model thoroughly. The modules in the model along with the parameters and outputs of each module are indicated in this step. Next, step 3 is "Verification Planning" (section 3.3.3). Based on the level of verification determined in step 1 and the structure of the simulation model identified in step 2, the process and the approach to be used can be decided in this step. Therefore, a plan to verify a simulation model can be made and the user can proceed to the next phase.

After the aforementioned steps are finished, the next phase is "Execution of Plan" (section 3.4). In this phase, the first action is step 4, "Perform tests" (section 3.4.1). This part of the framework is to perform the tests indicated in the plan. In case the model fails in a test(s), the following steps are required. Step 5, "Analyse problem" (section 3.4.2) is to understand the potential problem. The following actions to deal with the problem are addressed in step 6-1 (section 3.4.3) and 6-2 (section 3.4.4). When all of the steps are addressed, the user can proceed to the end of the framework. One thing to be mentioned is that, at the end of each step, there will be a check point for the user/developer to assess whether the step is completed.

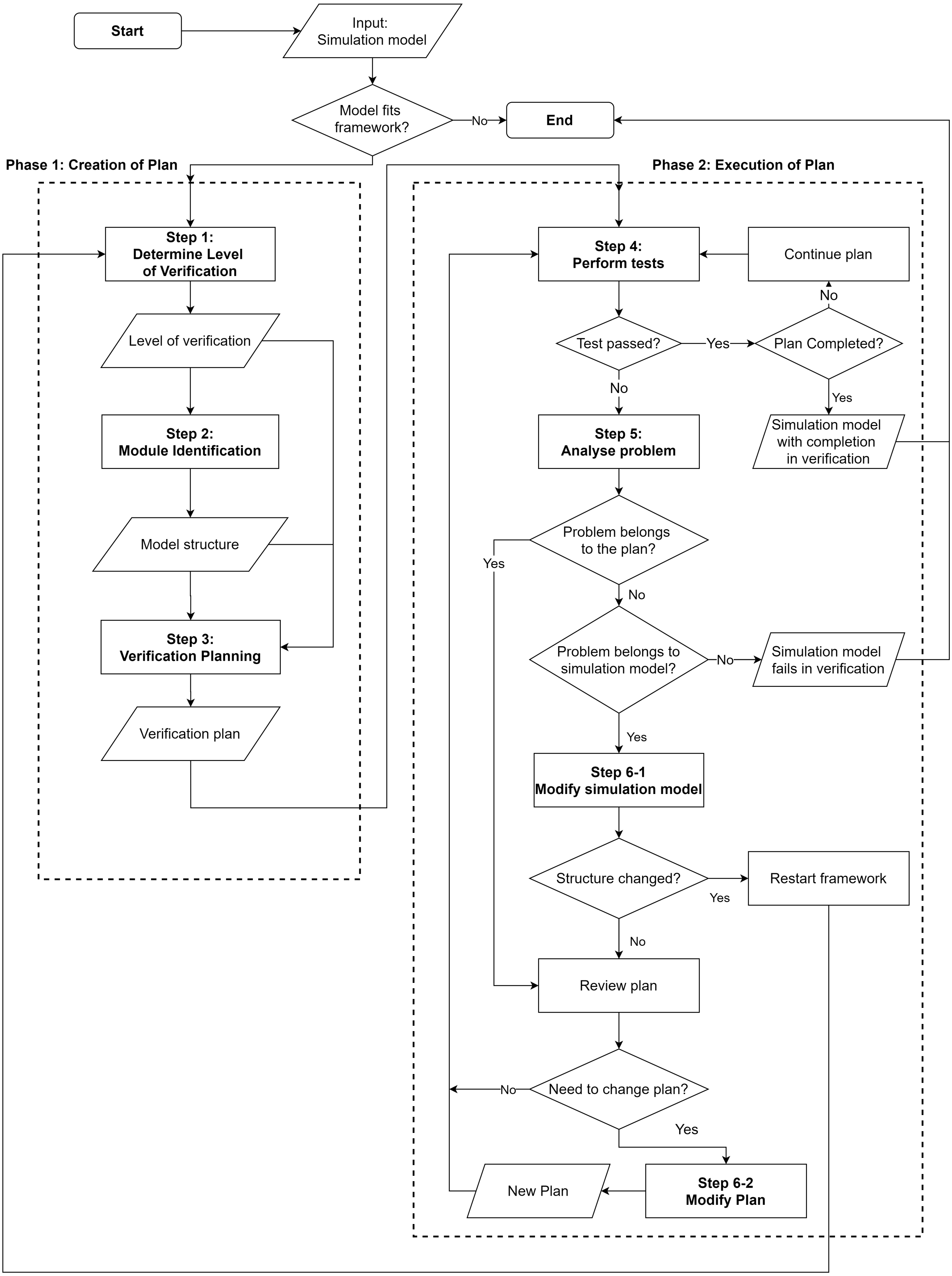

Figure 3.1: Structure of the proposed verification framework

3.2 START

Before practising the first step in the framework, users/developers must know whether they should use this framework or not. There are a few points to access whether the input simulation mode fits the framework. As the framework is designed for traffic simulation model, the elements and the decision-making process in the framework are assumed to be applied to a traffic simulation model. Therefore, the quality of the model along with the quality of the verification process cannot be guaranteed when applying the framework to other types of simulation model. Second, the framework is explicitly designed for verification of simulation model. The other processes in the development of a simulation model, such as implementation, validation or calibration, is not in the scope of this framework. Thus, it is not suggested to conduct other processes than verification. Third, in this framework, the conceptual model of the input simulation model is essential. Model verification is defined as a process to justify the implementation of the simulation model by comparing the simulation model with the conceptual model. Without a conceptual model, it will be unable to perform verification, as well as the proposed framework.

The three points mentioned above are summarized into the following questions for the users of the framework to perform a preliminary assessment on whether the framework can be used for their case.

Q3-1 Is the simulation model a traffic simulation model?

Q3-2 Is the simulation model required verification process in the development of a simulation model?

Q3-3 Is the conceptual model of the simulation model accessible?

If the answers to the three questions asked in this section are all "Yes", then it is able to proceed to step 1. If any of the answers are "No", it is suggested not to use the framework for the study, as the quality of the result is not guaranteed, and thus the study could be invalid.

3.3 PHASE 1: CREATION OF PLAN

After the input simulation model is checked, the user enter the the first phase of the verification framework "Creation of Plan" (figure 3.2). In this phase, the information of the input simulation model is first collected through the step 1,"Determine Level of Verification" and step 2, "Module Identification". The collected information is used to make a plan with tests of verification in Step 3, "Verification Planning". The preparation phase in the framework is completed when a plan is made.

3.3.1 Step 1: Determine Level of Verification

The first step of the verification framework is to determine the level of model verification. This step is used to collect information about the simulation model that identify the needed approach for verification. The content includes two subjects: identify the purpose of the simulation model and determine the level of model verification, as shown in the following figure 3.3.

When the level of verification is determined and processed, the user may continue to the question that is required to proceed to the next step. In the following subsections, the detail of the subjects is discussed.

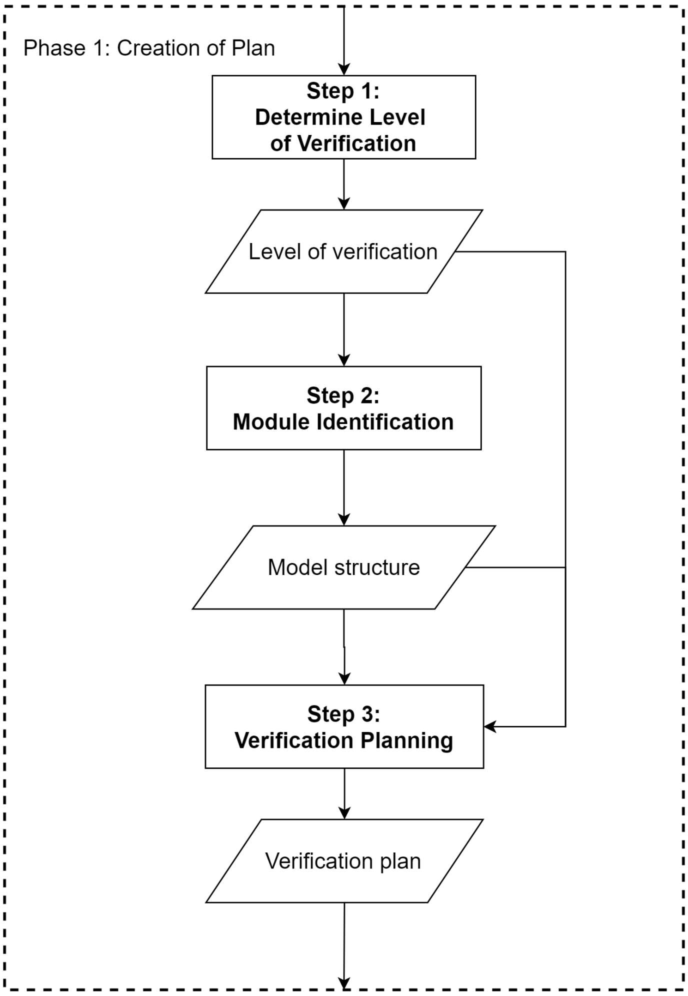

Figure 3.2: Phase 1: Creation of Plan

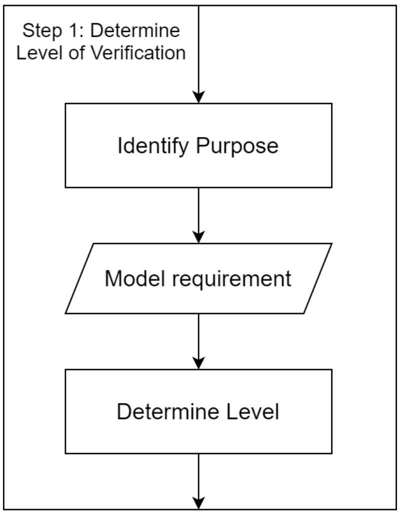

Figure 3.3: Step 1: Determine Level of Verification

Identify the purpose

To determine the level of verification, the first thing to understand is the purpose to develop the input simulation model. A simulation model is used to conduct an experiment where it is hard to practice a real-world experiment. Therefore, a simulation model is expected to represent (a part of) real world with some hypotheses and input data from real-world [Sargent, 2010]. The application of a simulation model can be broad or limited, depending on the purpose of the study as one of the principle from Balci [1998] said: "A simulation model is built with respect to the study objectives and its credibility is judged with respect to those objectives.". The main focus of recognising the purpose is to understand the aiming study area of the simulation model. The purpose, as a model requirement, has effect on the objective in model verification, thus influence the process and techniques to be applied. For example, if the input simulation model is only used in specific application, such as a study that only concerns a simulation within a certain study area or under certain conditions, it is not necessary to conduct sensitivity analysis or risk analysis on the model as long as the simulation model can work functionally within the specified study area. So, the question that should be answered regarding this subject is:

Q3-4 What is the purpose of the simulation model?

Determine the level

After the purpose of the input simulation model is recognized, the second thing to do is to determine the level of model verification. In the literature review, we concluded two level of model verification based on the two objectives to verify a simulation model. To ensure the basic functions of the simulation model within the study area, level 1 of verification is required. Based on the result of level 1, to understand the input-output relation in the simulation model and further find out the applicable range of the simulation model, level 2 of verification is required. Therefore, if the purpose to build the simulation model is to do an experiment within a certain study area, a verification of level 1 should be able to meet the requirement. However, if the input simulation model is meant to be a general tool for assessment, the applicable range is critical for the user/developer to know whether the model is able to produce credible results in a certain situation. Thus, a verification of level 2 should be aimed.

The two reasons to determine the level of model verification is that the level 1 of verification should comes prior than level 2 and the approaches used in different level of model verification are different. In level 1, as the verification process focus on the comparison between the conceptual model and simulation model, there will be more objective approaches that use statistical tests. In level 2, to understand the applicable range of the simulation model, subjective approaches that are applied without comparison to explore the model behavior are more often used [Sargent, 2010]. It is always suggested to complete both level of model verification. However, it might not be cost-efficient for all the case. Therefore, the question to be answered in this subject, also in this step, is:

Q3-5 What is the required level of verification according to the objective of model verification derived from the purpose of the model?

Check point for step 1

According to the answer to Q1, the target level of verification can be determined. The answer will affect the order of the process and the techniques to be used in the plan. Thus, the determined level, as an output of this step, will be the input to the later step of the framework as a support to make a complete plan. With the level of verification determined properly, the user can proceed to the next step. To check whether this step is completed, the following questions are asked to review this step:

Q3c-1 Is the purpose of the simulation model identified?

Q3c-2 Is the level of verification according determined?

If all the answers to the questions are 'Yes', the user may now proceed to the next step, "Module Identification".

3.3.2 Step 2: Module Identification

From step 1, the level of model verification is determined. Another task to be completed before planning the verification is to fully understand the structure of the input simulation model. This step, module identification, is dedicated to recognise the modules and the related elements inside a simulation model. The subjects including in this step is presented in figure 3.4.

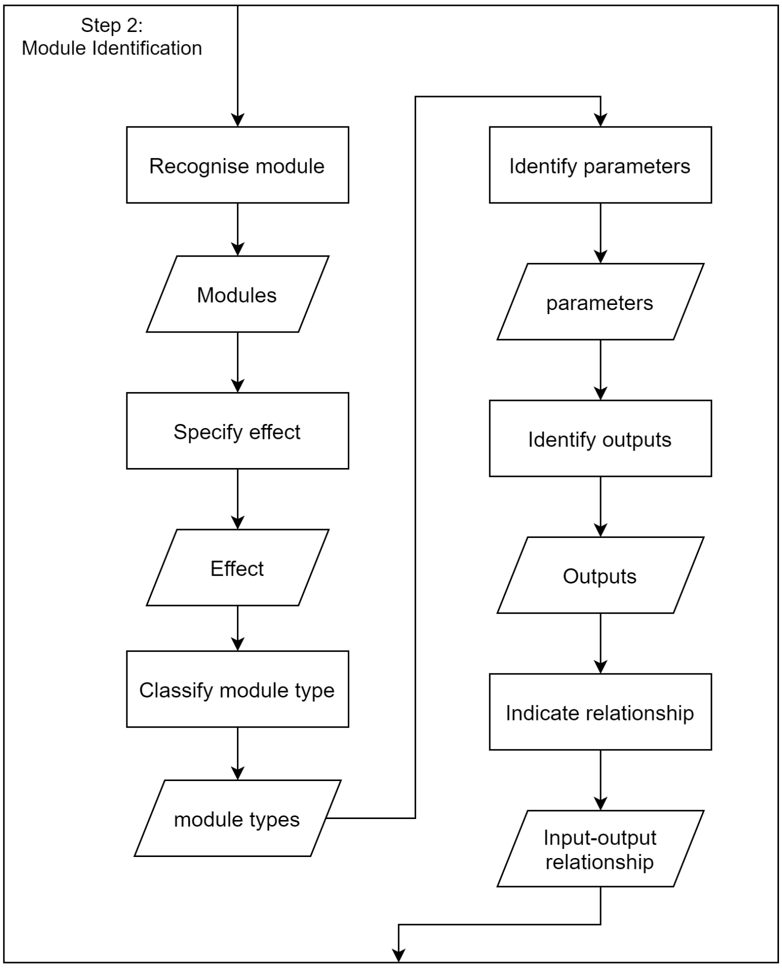

Figure 3.4: Step 2: Module Identification

From the literature study, we concluded the elements to be specified in a module are effect, module type, parameters, outputs (intermediate outputs and final outputs), and relationship. In the following paragraphs, each of the subjects to specify the elements in figure 3.4 is discussed.

Recognise module

As mentioned, a module is a device with a set of switch inputs and readout indicators [Parnas, 1972]. The first thing to do in module identification is to find out what are the modules included in the model. One way to do so is to acquire the specification of the simulation model from the model developer. The specification is a documentation containing the information of the functions and variables to be implemented in the simulation model. With the model specification, the user/developer only has to recognise the modules specified by the model developer.

As not all users can access the model specification, there is also other ways to complete this subject. If the source code of the input simulation model is accessible, reading through the code in the simulation model and recognising the paragraphs in the codes could be a way to do so, as each section of the code should have a certain function for the simulation model.

However, if the source code is not accessible, the action to recognize the modules would be challenging. The user can still assume the modules with the inputs and outputs of the model by categorising the related inputs and outputs together. But there would be uncertainty of the modules and the interaction within the model would be less clear. Therefore, it is suggested to applied the framework when the specification or the source code is available. No matter how the modules are identified, the question to be answered in this subject is:

Q3-6 What are the modules specified in the simulation model?

Specify effect

After the modules are recognised, the effect of each module needs to be specified. The effect means what the module do in the simulation model, for example handling the inputs from model user, processing the data, or generating a certain meaningful output. There are a few options to do so. In case the model specification from model developer is accessible, the user can use it as a reference to specify the effect of each module. If the model specification is not accessible, with the source code accessible, the user can grasp each section of the code to understand the mechanism of the module. If both model specification and source code is not accessible, it would be hard to specify the effect similarly. The question to be answer for each module in the subject is:

Q3-7 What is the effect of the module?

Classify module type

In the literature review, it is concluded that there are three types of module in a simulation model, which are mapping module, simulation module, and computerised conceptual model Sargent [2010]; Parnas [1972]. This sub-step is dedicated to categorise the modules according to their effects specified in the previous subject. The detail of each module type is addressed below:

Mapping module: This type of module provides redundant information. The value of this type of module is completely predictable from the current values of parameters or other readout functions. The mapping modules are provided as a notation to keep the programs clear and small [Parnas, 1972]. The value of the mapping module is predictable, so the verification process could be straightforward.

Simulation module: Simulation module is a type of module that is especially used to run a simulation. It could be the initialization of the simulation or other functions such as time-flow mechanism, random number generator, or random variate generators. The values of simulation modules might not be totally predictable, but they are expected to follow a rule or a certain distribution that determined by the user/developer. Therefore, the verification process of the simulation modules could also be straightforward.

Computerised Conceptual Model: This type of module is the core of a simulation model. Computerised conceptual model is the mathematical or logical representation of the problem entity that being implemented in the simulation environment using a programming language. The intermediate output of computerised conceptual model, could be predictable, but a calculation according to the logic and algorithm used in the model might be needed. The final output of this type of module is deemed as the outputs of the simulation model, and it is usually unpredictable.

According to the effect of the module specified in the last subject, it is possible to categorise the modules into these three types of module. The question to be asked for each module in this subject is:

Q3-8 Which type of module does the module belong to?

Identify parameters

After the previous subjects are completed, this subject is rather simple. The only thing the user has to do is to identify the parameters used in each module. The question for each module to be answered in the subject is:

Q3-9 What are the parameters used in the module?

Identify outputs

Outputs are the result generated from the module according to the certain set of parameter values. There could be intermediate outputs and the final outputs from the computerize conceptual models and they are different from model to model. Those indicators are the main point to decide which verification techniques to be used. Similar to the parameters, the outputs are identified based on the specified effect. So, the question for each module to be answered in the subject is:

Q3-10 What are the readout indicators generated from the module?

Indicate relationship

After the elements above are identified, the relationship between different modules need to be indicated in order to make a traceable path of model verification, since the output of a module could be an input to other modules. The question required to be answered in this subject is:

Q3-11 What are the relationships between the modules?

Check point for step 2

After finishing this step, the structure of the simulation model is understood. A module is a unit to be considered in this verification framework, and they are to be verified one by one in a verification. The reason to identify the modules is due to the fact that different modules required different verification process and techniques, according to their types, parameters, and outputs. The structure of the model is also a necessity to make a plan for verification. Table 3.1 shows an example of the result of this step as a documentation with the detail of the simulation model indicated.

Besides the table, a visualized structure could be helpful, especially when clarifying the relationship between the modules. The user can make one as the example presented in figure 3.5 when performing module identification. The consequence of each module is explicitly illustrated, as well as the type of the module, so it is possible to trace the cause of a certain result when verifying the simulation model.

Table 3.1: Example of documentation for Module Identification

Module	Effect	Module type	Parameters	Intermediate Output	Final Outputs	Relationship
Module 1	Initialiation	Simulation function	v0, x0, y0	v0, x0, y0	-	-
Module 2	Grouping of variable	Mapping function	v1,v2,v3	A = [v1, v2, v3]	-	-
Module 3	Decide acceleration	Computerized conceptual model	v, x, y, A	a(i)	a(i,t)	Input from: Module 1
...	...	...	...	...	...	...

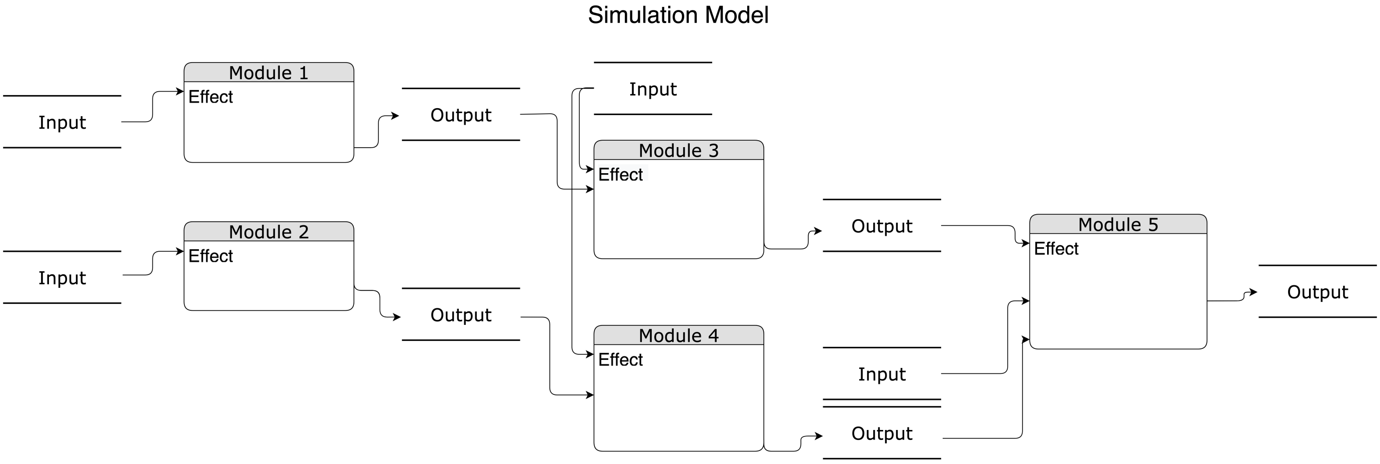

Figure 3.5: Visualization of module identification

The documentation, along with the product from the last subject, the level of verification, are to be the input to the next step to make a plan for verification. The detail is explained in the next section, planning of the verification process.

In this step, "Module Identification", the elements inside the input simulation model is thoroughly identified. Similar to step 1, all content should be covered before proceeding to the next step of the framework, therefore, a checklist (table 3.2) for the user of the framework is made to check whether a certain element is identified. If all the elements to be identified are confirmed with the check list, the user can proceed to the next step, "Verification Planning".

Table 3.2: Checklist for module identification

Element	Identified?
Modules	
Effects	
Module type	
Parameters	
Outputs	
Relationship	

3.3.3 Step 3: Verification Planning

In step 1, the level of model verification is determined, and in step 2, the elements in the model is identified. Following that, this step, "Verification Planning" is dedicated to make a plan for the verification process based on the information obtained in the previous subjects. Figure 3.6 presents the subjects to be completed for planning. It starts from "Determine process" to "Determined techniques" and ends up in "Create plan" by combining the products. In the following paragraphs, the subjects will be elaborated in detail.

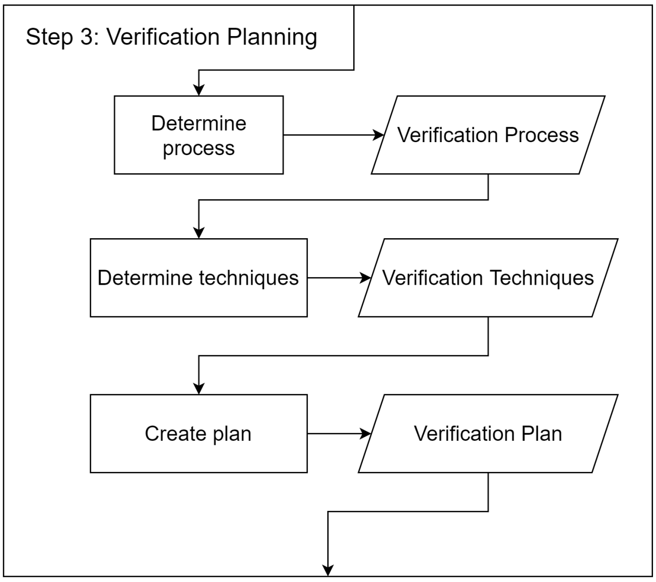

Figure 3.6: Step 3: Verification Planning

Determine Process

The process of verification is decided based on information obtained previously. According to the finding in the literature review there are three principles for the user of the framework to follow when planning the process: 1) level 1 of verification should come prior then level; 2) the planning for level 1 and level 2 are separately discussed, and 3) in level 1 of verification, module testing should come prior to model testing.

For principle 1, the target level of verification could be level 1 or level 2 as determined in the first step. The target level decides whether the user should plan for level 1 only or also level 2 of verification. As concluded in the literature, the basic functionality of the simulation model should be checked before performing other tests. Therefore, if the model requires level 2 of verification, level 1, assuring the basic functions of the simulation model, should comes prior to level 2, where the sensitivity of the model behaviour is investigated. To be more specific, both planning and execution of verification in level 1 should be fully finished before level 2.

For principle 2, as concluded from the previous studies, the two level of verification are quite different. The level 1 focus on comparing the simulation model and the conceptual model while level 2 aims to explore the model behavior using sensitivity tests. Therefore, for the planing of level 1, there exists a hierarchical order to assess the elements in the simulation one by one first, then examine the whole model. For level 2, as it is not possible to test all the combination of parameters of the model, it is important to select the critical parameters that influenced the performance of the model [Horiguchi and Kuwahara, 2002]. Therefore, the plan for level 2 should include a list indicating the critical parameters to be tested in verification. Because of the difference, the order to apply tests and the techniques to be used for the two levels of verification are discussed and performed separately.

For principle 3, as explained in the last paragraph, tests are performed module by module first in level 1 to assure the functionality of each module. After the function of each module is checked, tests on the whole model is to be performed. For the order to conduct tests on modules in level 1, it is suggested to follow the data flow of the simulation model according to information about the structure of the simulation model, avoiding unverified outputs being used as inputs to other module.

Based on the principles, the user can plan the process of verification. To be more specific, for level 1of verification, which module should be tested first can be decided. For level 2 of verification, which parameters to be tested can be selected.

Determine Techniques

After the process is determined, the second thing to be included in the plan is the approach of the verification. In this subsection, the selection of the techniques is discussed based on the measure introduced by Sargent [2010].

All of the techniques summarised in section 2.2.2 can be used for model verification. The major attribute affecting the selection of the techniques is whether the conceptual model is observable, where observable means that the data on a certain operational behavior to be compared can be collected from the conceptual model. Table 2.3 from the research of Sargent [2010] gives a classification of the techniques to be used based on the decision approach (subjective, objective) and observable of the conceptual model.

When the conceptual model is observable, which means that the output to be compared in the simulation model can also be collected from the conceptual model, the subjective approaches are "Comparison using graphical display" and "Explore model behavior", and the objective approach is "Comparison using statistic tests and procedures and procedures". "Comparison using graphical display" means comparing the data from the simulation model with the data from the conceptual model using graphical display. "Explore model behavior" means analysing the model behavior under different situations. "Comparison using statistic tests and procedures" is to use confidence intervals or hypothesis tests to make a comparison.

When the conceptual model is non-observable, the subjective approaches are "Explore model behaviour" and "Comparison to other model", and the only objective approach is "Comparison to other models using statistical tests". However, it is found in the literature review that the approaches of comparison to other models could be challenging for simulation models of traffic modes other than vehicle. It is not only because it could be hard to find another simulation model describing the same behavior of cyclist with a small number of researches, but also because it is hard to tell whether the comparison is valid due to the lack of critical values as reference values of the behaviour of other modes. One thing to be noticed is that, it is preferable to use statistical tests for the comparisons because these allow for objective decisions [Sargent, 2010].

The classification of different approach is the principle to select different techniques for verification. As aforementioned, the planing for level 1 and level 2 should be separated. Therefore, we classified the techniques according to the level of verification in the literature review. In the following paragraphs, the selection of the techniques for different level of verification is discussed.

TECHNIQUES FOR LEVEL 1 OF VERIFICATION: In table 2.4, the techniques from previous literature for level 1 of verification are summarised. To assure the functionality of the simulation model, the techniques to be used in level 1 are to check the consistency and correctness of the module, and make sure the simulation model is capable of reproducing the results of the conceptual model or capable of performing intended application. In terms of testing the module, the techniques to be used are selected based on the type of the module. For the those modules with predictable value, such as *mapping modules* and *simulation modules*, the technique to be applied is basically a consistency check. For *computerised conceptual model*, consistency check is also applied for the intermediate output. A face validation with the calculation of the algorithm might be needed when there is difficulty to predict the intermediate output directly. As for the whole model, to test the final outputs of the computerised conceptual model, statistical tests is recommended. Techniques such as animation and graphical comparison are optional depending on the model behavior. If the simulation model is built to represent a dynamic system, animation is then suggested.

One thing to be mentioned is that the choice of statistical test depends on the outputs of the model. It could be parametric or non-parametric. In the study of Nayak and Hazra [2011], they introduced five example questions with related schemes to demonstrate how to select a suitable statistical test. In here, the author summarised the most common ones to be used when doing a comparison between conceptual model and simulation model:

- Is there a difference between groups that are unpaired?
  This question is used when comparing two groups of data and they are not related to each other. For numerical data, if the two datasets follow normal distribution, parametric tests is to be applied. If the datasets do not follows normal distribution it is suggested to use non-parametric tests. In terms of categorical data, Chi-square test is suggested. The specific tests are indicated in the figure 3.7.
- Is there a difference between groups which are paired?
  This question is asked when comparing two groups of data and they are related to each other. According to Nayak and Hazra [2011], "pairing signifies that data sets are derived by repeated measurements (e.g. before-after measurements or multiple measurements across time) on the same set of subjects". This is one of the common situations to test the model as there could be a subject to be compared with one measurement from simulation model and the other from conceptual model. Similarly, the specific tests are indicated in the figure 3.8.

A brief explanation on how to selected statistical tests is therefore introduced. However, there could be situations which are not considered by the two example questions. In that case, the user of the framework can refer to the other example questions in the study of Nayak and Hazra [2011], or other researches, such as *"How to select appropriate statistical test in scientific articles?"* from Trajkovski [2016] for further information on the selection of statistical tests.

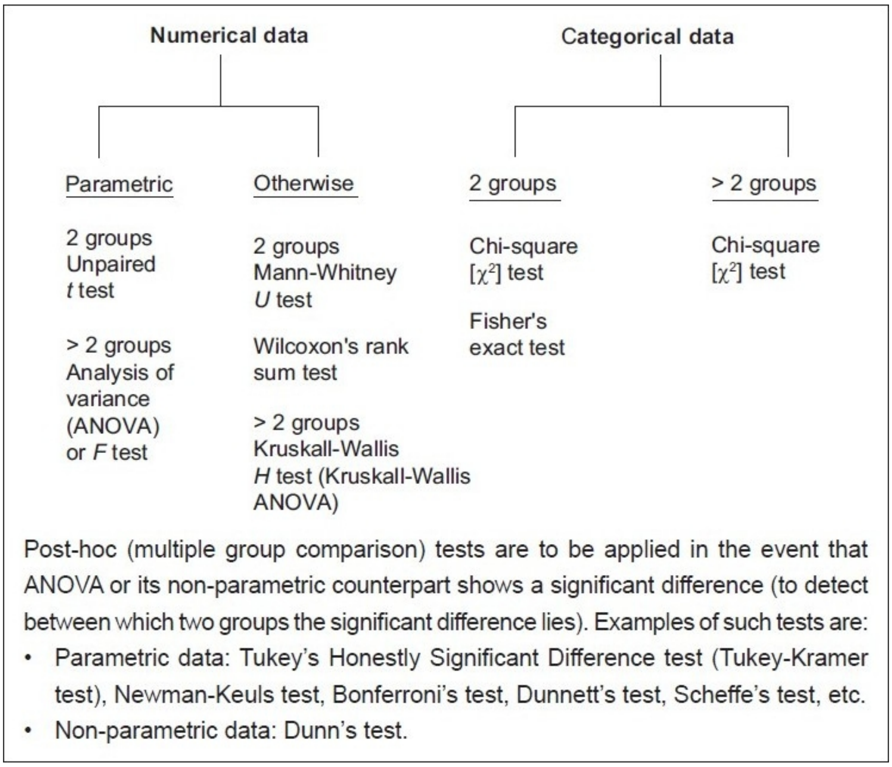

Figure 3.7: Scheme for the first question from Nayak and Hazra [2011]

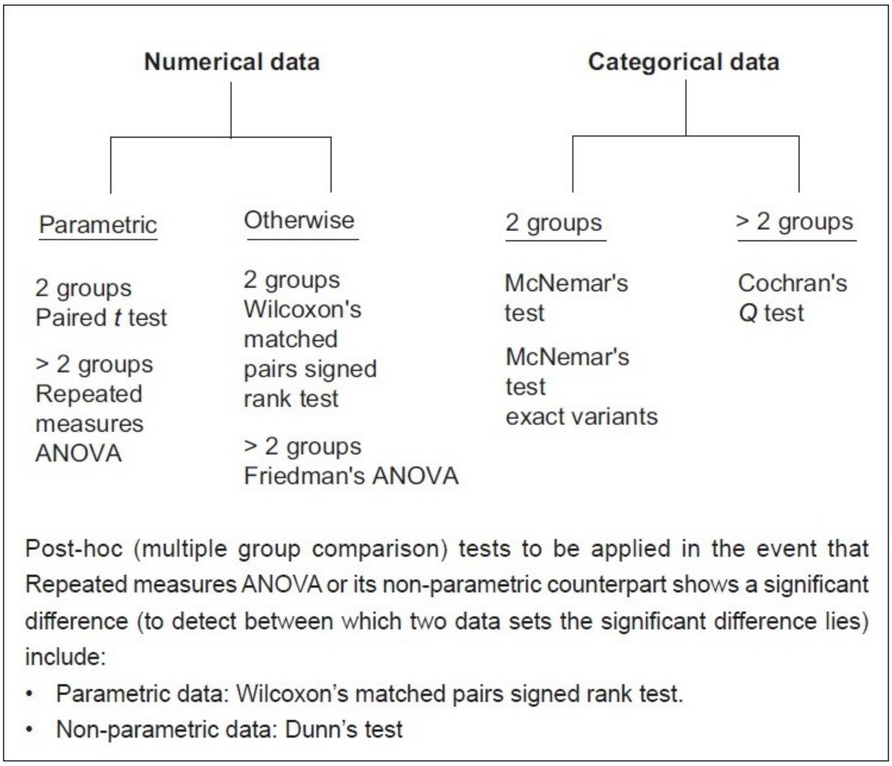

Figure 3.8: Scheme for the second question from Nayak and Hazra [2011]

TECHNIQUES FOR LEVEL 2 OF VERIFICATION: Table 2.4 also displays the techniques from previous literature for level 2 of verification. To decide the applicable range of the simulation model, the main task in level 2 of verification is sensitivity analysis. Various sets of experimental conditions from the study area of the model's intended application should be included for comparison and exploring model behavior. Furthermore, to understand the limit of the simulating model, special (extreme) experimental conditions also need to be considered, such as special input testing.

Unlike level 1 of verification, the techniques for level 2 are not specific for a module. Instead, it is about verifying the whole model. Because of that, the sensitivity analysis is performed mainly on the intermediate output and the final output of the computerised conceptual model. Other subjective approaches such as animation and graphical comparison are used to assist the user to understand the model behavior in different conditions. Functional testing, as risk analysis, can be used when the robustness of the simulation model is required.

In terms of the design of a sensitivity analysis, it includes combinations of parameters that are to be varied. A design should also indicate how many parameters need to vary at a time in order to to study the correlations. Therefore, the parameters to vary and the according results are to be selected first. These parameters should be selected based on their influence to the model, as the critical parameters motioned by Horiguchi and Kuwahara [2002]. There are two types of measurement for sensitivity analysis: local and global. Local sensitivity analysis is a one-factor-at-a-time technique that analyzes the impact of one parameter at a time, keeping the other parameters fixed, while global sensitivity analysis is often implemented using Monte Carlo techniques to explore the design space [Morio, 2011]. Therefore, local sensitivity analysis can be used to determined the applicable range of each parameters, its effect on the model, and also the suggested value(s) when the performance of a model is measurable. On the other hand, global sensitivity analysis can be used to decide the robustness of the model in different situations. However, in that case, the contribution of each parameter to a certain result might be unknown, so from the author's perspective, it is suggested to perform a local sensitivity analysis before a global one if the robustness is to be studied.

Nevertheless, the requirements of a sensitivity analysis could vary with different models. In case the requirement is not included in the explanation by the author, the user can further refer to the studies from Saltelli et al. [2004], Saltelli et al. [2008], or Komkov et al. [1986] for comprehensive discussion on sensitivity analysis.

Create plan

After the process and techniques are determined, a whole plan including level 1 of verification and level 2 of verification (if needed) can be made. The plan could be in any format as long as it indicates the order to verify the module/model, and the techniques to be used for different level of verification. A visualized plan could also help the user of the framework to follow when conducting verification. However, a documentation with table is more recommended. Table 3.3 shows an example of detailed plan with indication of level of verification, module, module type, techniques, and whether it is tested/checked. This will create a clear overview, and therefore easy for the user to process the next phase, "Execution of Plan".

Table 3.3: Documentation for Verification Planning

Level 1							
Order	**Module**	**Module Type**	**Parameter**	**Intermediate Output**	**Final Ouput**	**Techniques**	**Test Passed**
				Module Testing			
1	Module 2	Mapping function	v1,v2,v3	A = [v1, v2, v3]	-	Consistency Check	
2	Module 1	Simulation function	v0, x0, y0	v0, x0, y0	-	Consistency Check	
3	Module 3	Computerized conceptual model	v, x, y, A	a(i)		Consistency Check Face Validation	
...	...	...	...	...	...	...	...
				Model Testing			
4	Module 3	Computerized conceptual model	v, x, y, A		a(i,t)	Statistical Test Graphical Comparison	

Level 2				
Order	**Parameter**	**Module of Parameter**	**Correlations**	**Checked**
		Local Sensitivity Analysis		
1	v1	Module 2	a(i), a(i,t)	
2	v0	Module 1	a(i), a(i,t)	
...	...	...		...

3.3.4 Check point for step 3

In this step, the planning for model verification is discussed. The process and the techniques to be applied for verification are included in the plan. All of these techniques should be applicable at the specific module in the simulation model or at the model itself and meet the objective in the certain level of verification. The following questions are asked to assess the completeness of this step:

Q3c-3 Is the process to conduct of verification determined?

Q3c-4 Are the orders to perform tests in each level of verification identified?

Q3c-5 Are the techniques for the level 1 of verification selected?

Q3c-6 Are the techniques for the level 2 of verification (if existed) selected?

Q3c-7 Is the plan completed?

If all of the answers are yes, the user can proceed to the next phase, "Execution of Plan".

3.4 PHASE 2: EXECUTION OF PLAN

This section is dedicated to the execution of the plan that created in the previous phase. The main goal in this phase is to execute and complete the verification plan made in step 3. Ideally, by performing the tests with the techniques selected in the plan step by step according to the determined process, the simulation model can be systematically verified. However, there is a chance that a certain module(s) fails in the selected test(s). In that case, other actions are needed.

In figure 3.9, the execution steps are merged into an iterative process. The process starts with "Perform techniques". If the module/model pass the test, then the user can continue the next test to be perform until the plan is completely executed. In case the module/model fails to pass the test, the user needs perform the following steps to identify and solve the problem: "Analyse problem", "Modify simulation model" (if required), and "Modify plan"(if required). In the following paragraphs, the steps in figure 3.9 will be elaborated in detail.

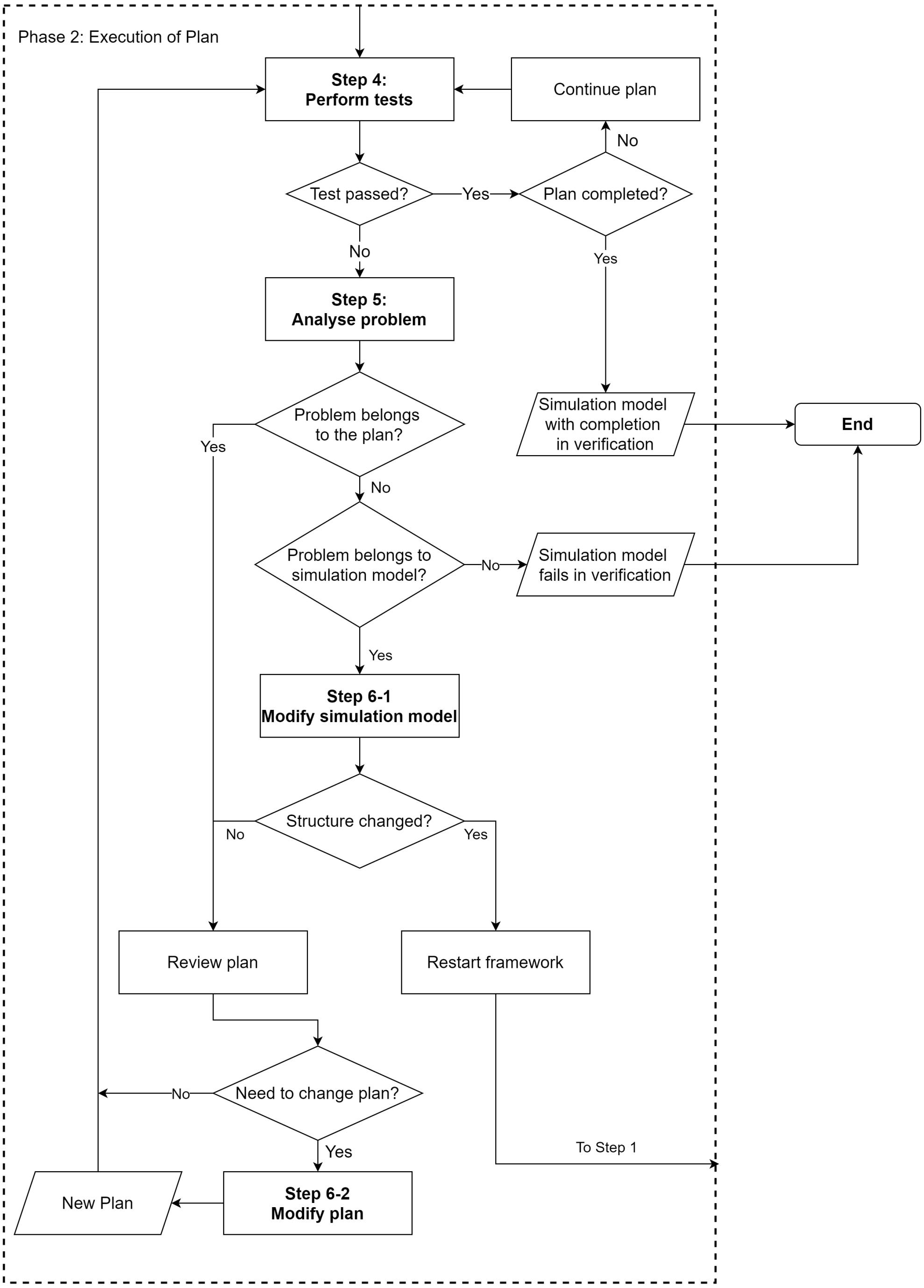

Figure 3.9: Phase 2: Execution of Plan

3.4.1 Step 4: Perform techniques

The techniques selected to verify the input simulation model must be performed carefully. For subjective approach, simulation model verification requires independence to prevent user's bias. According to the study of Balci [1998], he said "the model developer with the most knowledge of the model may be the least independent when it comes to testing". To prevent such bias, a double check or consultancy from a third party is plausible.

For objective approach Balci [1998] explained three types of errors, type I, type II and type III, to be prevented when conducting verification. A type I error occurs when the outputs of the module/model are rejected when in fact they are creditable. A type II error is committed when invalid module/model outputs are accepted as if they are valid. A type III error occurs when the wrong problem is solved, resulting from the fact that the problem formulated does not represent the actual problem.

The user must focus on minimizing these risks as much as possible when performing the techniques, since the cost of model development could greatly increase due to these errors. Also, the consequences of type II and type III errors could be serious when critical decisions are made on the output with the errors [Balci, 1998].

If the module pass the test, the user can proceed to the next test indicated in the plan until the plan is completed. When a certain module or the model fails in a test(s), the user of the framework needs to find out the problem and fix it if possible. The first action to discover the reason why it fails is to analyse the problem, which is discussed in the next subsection.

3.4.2 Step 5: Analyse problem

This step, Analyse problem, is dedicated to find out the reason why the certain module or simulation model fails in the test. When the source code is available, the logic, syntax, and functions related to the problem are to be examined. When the source code is not available, there should be an inspection on the functions that assumed existing in the model. However, it could be difficult to fix the problem in the later steps without the access to the source code even the problem is identified.

There are quite many techniques that could be used to inspect the simulation model as show in figure 3.4. The techniques are summarized from the verification and validation tests mentioned in previous studies [Kleijnen, 1995; Balci, 1998; Sargent, 2010] that are more in software development perspective. Depending on the case, it could take a long time and require many techniques to figure out the problem. Also the following question is used to analyse the problem:

Q3-12 Why does the module/model fail in the selected test?

Table 3.4: Techniques to be used in inspection

Inspection Techniques	
Bottom-Up Testing	Reviews
Cause-Effect Graphing	Semantic Analysis
Control Analysis	Structural Analysis
Data Flow Testing	Sub model/Module Testing
Fault Failure Analyst	Symbolic Debugging
Functional Testing	Symbolic Evaluation
Inspections	Syntax Analysis
Interface Testing	Top-Down Testing
Logical Deduction	Turing Tests

Before looking into the development of the simulation model, the first thing to inspect is the plan made in the first phase. There is a chance that the selected test is not suitable for the module/model to be tested due to to conditions like misunderstanding of the parameters used or the readout indicators during "Model Identification". In that case, the problem can be solved simply by performing step "Modify plan". Therefore, the second question to be answered in this step is:

Q3-13 Does the problem belong to the plan created in the previous phase?

In the development of a simulation model, the problem could occur in modeling (from system to conceptual model) or in implementation (from conceptual model to simulation model). As the verification is concerning the justification of the implementation, the problem occurs in modeling is assumed fixed before implementation. If the problem is found out related to modeling process, which is related to the justification of conceptual model instead of simulation model, a further inspection to the problem regarding conceptual model is needed. In that case, the framework would be no longer feasible to the input simulation model, thus come to "End" of the framework, where there should be no action anymore. (In case a modification or a validation of the conceptual model is required, the user can refer to the study of Sargent [2010] for further information about conceptual model validation.) Therefore, the third question to be answered in this step is:

Q3-14 Does the problem belong to the simulation model?

If the problem is found related to the implementation of the simulation model, the user may proceed to the next step for fixing the problem, "Modify simulation model".

3.4.3 Step 6-1: Modify simulation model

Once the problem is analysed and concluded that the problem is within the implementation of simulation model, a modification should be done if possible. If the source code is not accessible, the user of the framework should report the problem to the model developer(s) so that they can modify the model and fix the problem. The modification is a trial-and-error process of debugging according to the problem analysed in the previous sub-step. Basically, there are two types of error in implementation, syntax error and semantic errors. Syntax error is relative easy to be solved, as most of the syntax error can be detected by the compiler. For semantic error, the user/developer has to go through the specification along with the code to catch the problem.

Before proceeding to the next sub-step, the user needs to assess the scale of the modification. If the problem occurs in a module like mapping module or simulation module, which has little effect on the structure of the simulation model, the verification plan made previously could be used with a review and modification. However, if the problem is found out related to the computerised conceptual model, or other modules that have intensive connection to other modules that fixing it might change the structure of the simulation model, a restart of the framework is recommended. Therefore, the following question is asked in this step:

Q3-15 Is the structure of the model changed due to the modification of the simulation model?

Based on the answer to the question, the user may proceed to subject "Review plan" or "Restart framework" in the framework.

REVIEW PLAN: The user is suggested to make a review on the plan whenever a modification is performed since there might be a change in parameters, outputs or relationship between modules, which would lead to different techniques to be used or a different process to follow. To make a review of the plan, the following questions are helpful when assessing whether the plan still fits the model.

Q3-16 Is the determined process feasible to the changed model?

Q3-17 Are the selected techniques suitable for the changed module?

If both of the answer is yes, the plan made is feasible to the changed model. The user can start to execute the plan again and continue to complete the plan. If any of the answers toward the questions above is ”No”, the user needs to modify the plan made previously, which will lead to another step for fixing the problem, ”Modify plan”, which will be discussed in next subsection.

RESTART FRAMEWORK: In case the modification affects more than one modules, which leads to a change in the structure of the simulation model, it is suggested to revisit the framework from the beginning as the scale of the change is large enough to affect the plan made previously. Therefore, the user should restart from the first step (section 3.3.1) of the framework.

3.4.4 Step 6-2: Modify Plan

Basically, to modify a plan, the user has to go through each step of the framework again for the certain modules. Those modified modules needs to carry out ”module identification” first to check whether there is change in specified modules, effect, module type, parameters, outputs, and its relationship with other modules. The techniques and process needs to be determined again based on the classification mentioned in section 3.3.3. After that, a modified plan is made, and the user can resume to the execution of the plan again.

3.4.5 Check point for phase 2

Unlike phase 1, the steps in phase 2 are a series of continue actions to execute the plan. Therefore, they are to be assessed together in this subsection. When all of the plan is completed, the user can proceed to the end of this phase, also obtaining the product of the framework, a simulation model completed its verification process. In this phase, the uncertainty of the verification is introduced and considered in the framework. The end of this step represent the complete of plan created in phase 1 as well as the complete of the model verification. The techniques indicated in the plan should be performed thoroughly and the execution should be reviewed.

In the end, all tasks on the list of the plan should be checked in table 3.3, including the comment on the problem met and whether the problem is solved. This will create a clear overview of the completeness of verification. When the input simulation model passed all the test in the plan, it becomes a verified simulation model ready for the next step of model development, such as calibration and validation, and therefore come to the end of the framework.

4 FRAMEWORK APPLICATION: VERIFICATION OF THE QUEUE FORMATION SIMULATION MODEL

This chapter is dedicated to the application of the framework in a real case. First, the choice of the queue formation model is discussed, followed why this particular case is suited for the methodology application. Then, the framework is applied in the case of the queue formation simulation model with the steps discussed in the previous chapter. To elaborate on the application of the framework, each step is discussed on two levels: the framework feasibility on the case and case-related content. The content that is related to the case is presented between horizontal lines, starting with a single line and ending with double lines.

4.1 CASE STUDY: THE CYCLIST QUEUE FORMATION SIMULATION MODEL

The queue formation model is a microscopic cyclist operational model that describes the queue formation process of cyclists at a signalised intersection. Gavriilidou et al. [2019b] introduced a two-layer framework of discrete choice model to represent the operational decision-making process at the mental layer and the physical layer. The main purpose of the study is to create a simulation tool to assess the cycling infrastructure.

As a response, the Active Mode Lab (Department of Transport & Planning under Faculty of Civil Engineering and Geosciences) in TU Delft is currently developing a simulation model based on the queue formation model. The simulation model has been implemented in Matlab to test its feasibility, and verification process to assess its basic functionality remains to be done. Therefore, the queue formation simulation model is selected to test the applicability of the framework to a cyclist simulation model due to the fact that it meets the criteria as a cyclist simulation model, as a simulation model required verification, and most important, it is accessible for the author to conduct this research.

4.2 APPLICATION STEP: START

As the framework introduces, first, the start procedure should be executed. In the following paragraph, the questions to assess whether the input module fit the framework and corresponding answers are presented.

Q3-1 Is the simulation model a traffic simulation model?

Ans: Yes. The cyclist queue formation model is a simulation model that represents the cyclist traffic behaviour. Thus it is a traffic simulation model.

Q3-2 Is the simulation model required verification process in the development of a simulation model?

Ans: Yes. The queue formation simulation model has been implemented in a simulation environment from the specification of its conceptual model. The current facing procedure in simulation model development is model verification.

Q3-3 Is the conceptual model of the simulation model available?

Ans: Yes. The conceptual model made by Gavriilidou et al. [2019b] is accessible from the Active Mode Lab of TU Delft. The input data and the corresponding outputs are also available.

All answers show that it is positive to proceed with the methodology. Therefore, the cyclist queue formation simulation model is suited as a case for the framework application. In the following sub-sections, the results of the framework application are discussed.

4.3 APPLICATION STEP 1: DETERMINE LEVEL OF VERIFICATION

This section is dedicated to the application of the first step, "Determine Level of Verification", of the framework. This is also the beginning of the first phase, "Creation of Plan" in the framework. As described in the framework, there are two subjects to be recognised in this step.

Identify Purpose

In this sub-step, the question regarding the purpose of the input simulation model is answered.

Q3-4 What is the purpose of the simulation model?

Ans: The queue formation simulation model is currently being tested on its basic functions in a simulation environment. Therefore, the aiming use of the queue formation simulation model is still in the specific study area where the conceptual model is proposed.

Determine Level

From the purpose of the study, the question related to the level of verification can be answered.

Q3-5 What is the required level of verification according to the objective of model verification derived from the purpose of the model?

Ans: As the answer to Q3-4 described, the application of the queue formation model is still in the study area where the conceptual model was built on. Because of that, the verification of the queue formation simulation model focuses on the basic functions of the simulation model, assessing its ability to reproduce the results of its conceptual model and explore its model behavior to see whether it performs as intended. Based on the purpose, we can conclude that the cyclist queue formation simulation model requires a verification of level 1.

As the framework described in the previous chapter, both subjects are covered in the application, enabling the possibility to continue to the next step, 'Module Identification'. This will be discussed in the following sub-section.

4.4 APPLICATION STEP 2: MODULE IDENTIFICATION

For the application of the second step, "Module Identification", the elements in the cyclist queue formation simulation model are specified. The questions in each sub-step are used to determine the content of the respective subjects.

Recognise module

The first element to be specified in this step is the module. The question regarding the modules in the simulation model is discussed.

Q3-6 What are the modules specified in the simulation model?

Ans: From the understanding of the simulation model, seven modules are specified in the cyclist queue formation simulation model, which are: Mental Choice Set Construction, Physical Choice Set Construction, Simulation Parameter Defining, Mental Layer Discrete Choice Model (DCM Mental), Physical Layer Discrete Choice Model (DCM Physical), Simulation Running, and Visualization.

Specify effect

Based on the modules identified in the first sub-step, the effect of each module can be described. In this sub-step, the question regarding the effect of each module is discussed.

Q3-7 What is the effect of the module?

Ans: The function of each module is specified in the following table:

Table 4.1: The effects specified from each module of the queue formation simulation model

Module	Effect
Mental Choice Set Construction	Define the mental layer choice set
Physical Choice Set Construction	Define the physical layer choice set
Simulation Parameter Defining	Initialisation of the parameters used for simulation
DCM Mental	The discrete choice model that describe the preference to the stop position of a cyclist
DCM Physical	The discrete choice model that describe the preference to the change in steering angle and speed of a cyclist
Simulation Running	1. Run the simulation according to the simulation setting. 2. Generat cyclists 3. Handle the input and output to and from the DCM Mental and DCM Physical
Visualisation	Display the animation of the results.

Classify module type

According to the understanding of each module, it is possible to classify each module into the three types of module mentioned in the framework. The three types of module are mapping module, simulation module, and computerised conceptual model. The question regarding the type of each module is discussed.

Q3-8 Which type of module does the module belong to?

Ans: The module type of each module is indicated in the following table:

Table 4.2: Classification of the module type for each module in the queue formation simulation model. (The '...' is the omission of the information mentioned in the previous table(s). This applies to all of the following tables.)

Module	Effect	Module type
Mental Choice Set Construction	...	Mapping module
Physical Choice Set Construction	...	Mapping module
Simulation Parameter Defining	...	Simulation module
DCM Mental	...	Computerized conceptual model
DCM Physical	...	Computerized conceptual model
Simulation Running	...	Simulation module
Visualisation	...	Mapping module

Identify parameters

When the effects and the module types are specified, the next sub-step is to identify the parameters used in each module. The question regarding the parameters is discussed.

Q3-9 What are the parameters used in the module/model?

Ans: The parameters used in each module are presented in the following table. Due to the fact that there is a large number of parameters, it is hard to show all of them. Also, using the symbols only could be confusing. Therefore, the table only display the main parameters with short description. The full version of the parameters used in each module can be found in Appendix A (table A.1).

Table 4.3: The parameters used in each module of the queue formation simulation model that identified in this step.

Module	Effect	Module type	Parameters
Mental Choice Set Construction	…	…	Environment Parameter (Xmin, Xmax, Ymin, Ymax)
Physical Choice Set Construction	…	…	Physical Parameter (αmax, θmax)
Simulation Parameter Defining	…	…	Simulation Setting (Timestep, cycle time, Vmax, Vdesire)
DCM Mental	…	…	Choice set, Attributes, Coefficients (Xg, Yg, agentBike, Betas)
DCM Physical	…	…	Choice set, Attributes, Coefficients (Xg, Yg, agentBike, Betas)
Simulation Running	…	…	Timestep, Inflow, Cyclist state (Tk, Newflow, agnetBike)
Visualisation	…	…	Simulation result (trajectory_x, trajectory_y)

Identify outputs

After the parameters are identified, the next element to be specified is the output of each module. One thing to be mentioned is that there are two types of output from a computerised conceptual model, which are intermediate output and final output. In this sub-step, both types of output from each module are discussed according to the question.

Q3-10 What are the readout indicators made from the module?

Ans: The outputs from each module are presented in the following table with the symbols along with short description. The intermediate output and final outputs are separately displayed.

Table 4.4: The intermediate outputs (if exist) and the final outputs (if exist) of each module that identified in this step.

Module	Effect	Module type	Parameters	Intermediate Outputs	Final Outputs
Mental Choice Set Construction	...	...	...	Mental Choice Set (Xg, Yg)	-
Physical Choice Set Construction	...	...	...	Physical Choice Set (Fg, Vg)	-
Simulation Parameter Defining	...	...	...	-	-
DCM Mental	...	...	...	Preference to a stop position for a cyclist (P(i))	Probability distribution of the cyclists stop position (P*(i))
DCM Physical	...	...	...	Preference to a state of a cyclist (P(j))	Trajectories of the cyclists (trajectory_x, trajectory_y)
Simulation Running	...	...	...	New cyclists (agnetBike)	-
Visualisation	...	...	...	Animation	-

Indicate relationship

The last thing to be specified in this step is the relationship between different modules. In this sub-step, the question regarding the relationship is discussed.

Q3-11 What are the relationships between the modules?

Ans: The answer to the question can be obtained from the parameters and the outputs of the module. The relationship is shown in the following table.

Table 4.5: The relationship from one module to another module that indicated in this step.

Module	Effect	Module type	Parameters	Intermediate Outputs	Final Outputs	Relationship
Mental Choice Set Construction	...	...	...	...	...	Output to DCM Mental
Physical Choice Set Construction	...	...	...	...	...	Output to DCM Physical
Simulation Parameter Defining	...	...	...	...	...	
DCM Mental	...	...	...	...	...	Output to DCM Physical (Final result) output to Visulization
DCM Physical	...	...	...	...	...	(Final result) output to Visulization
Simulation Running	...	...	...	...	...	Output to DCM Mental and DCM Physical
Visualisation	...	...	...	...	...	

After finishing the last sub-step, the elements to be specified in this step are all completed. The completed version of the table can be found in the Appendix (table A.2). The visualised structure is also made as a support to understand the model in the Appendix (figure A.1). With all the necessary elements in the modules specified, the next step of the framework can be applied, which is the planning of verification and it is discussed in the next sub-section.

4.5 APPLICATION STEP 3: VERIFICATION PLANNING

This sub-section is dedicated to the execution of the step "Verification Planning" of the framework. In the following paragraphs, the content of making a plan for verification is discussed as the framework described.

Determine Process

The first sub-step in this step is to determined the process of verification. This sub-step basically decides the order to perform test in both level of verification.

Since it is determined to conduct level 1 of verification, the principle to decide the process in level 1 is to make sure that the modules are tested first before model, and orders within the module is based on the data flow of the simulation model. Based on this principle, the result of this sub-step is presented in the following table.

Table 4.6: Determine process in level 1. (The cell with "..." implies the omitted information that is mentioned in the previous step, and the cell with "-" means there is no information in the cell. This is also applied to the following tables)

Level 1					
Order	**Module**	**Module type**	**Parameters**	**Intermediate Outputs**	**Final Outputs**
Module Testing					
1	Mental Choice Set Construction	Mapping module	...	...	-
2	Physical Choice Set Construction	Mapping module	...	...	-
3	Simulation Parameter Defining	Simulation module	...	...	-
4	Simulation Running	Simulation module	...	...	-
5	DCM Mental	Computerized conceptual model	...	...	-
6	DCM Physical	Computerized conceptual model	...	...	-
7	Visualisation	Mapping module	...	...	-
Model Testing					
8	DCM Mental	...	...	-	...
9	DCM Physical	...	...	-	...

Determine Techniques

After the process is determined, the last thing to be decided in phase 1 is the techniques to be used in verification. Following the approach introduced in the framework, the techniques to verify each module and the model in level 1 of verification can be determined. The selection of techniques is discussed in the following paragraph.

For level 1 of verification, the techniques for module testing are first determined. In terms of the mapping modules (Mental Choice Set Construction, Physical Choice Set Construction, and Visualisation) and simulation modules (Simulation Parameter Defining and Simulation Running), due to the fact that their outputs are predictable, *consistency check* for the correctness of the implementation is selected. For the computerised conceptual model (DCM Mental, and DCM Physical), besides *consistency check* on the correctness of the code, a *face validation* of calculation according to the discrete choice model is selected.

As for model testing, *statistical tests* is chosen for DCM Mental with the outputs of interest observable in the conceptual model. In the study of Gavriilidou et al. [2019b], the conceptual model validation of DCM Mental and DCM Physical were done with a face validation on the graphical displays of the probability distribution of both modules. The graphical displays showed the probability of cyclist stop positions for DCM Mental and the probability of the change in speed and steering angle for DCM Physical. This face validation is also selected for DCM Mental to have a quick assessment before conducting statistical tests.

Unlike DCM Mental, the output of interest for DCM Physical is not observable in the conceptual model since there is no data related to the trajectories in the conceptual model. Because of that, the approach to verify this module in model testing is to explore the model behaviour. In terms of the technique, *graphical comparison* becomes a credible choice since the trajectory data usually presents in image. Besides, a face validation similar to the conceptual model validation mentioned in the previous paragraph is also selected to have a quick assessment of the model.

Table 4.7: Determine techniques in level 1

Level 1								
Order	**Module**	**Module type**	**Parameters**	**Intermediate Outputs**	**Final Outputs**	**Techniques**	**Test Passed (Y/N)**	**Other**
Module Testing								
1	Mental Choice Set Construction	Mapping module	…	…	-	Consistency Check		
2	Physical Choice Set Construction	Mapping module	…	…	-	Consistency Check		
3	Simulation Parameter Defining	Simulation module	…	…	-	Consistency Check		
4	Simulation Running	Simulation module	…	…	-	Consistency Check		
5	DCM Mental	Computerized conceptual model	…	…	-	Consistency Check Face Validation		
6	DCM Physical	Computerized conceptual model	…	…	-	Consistency Check Face Validation		
7	Visualisation	Mapping module	…	…	-	Consistency Check		
Model Testing								
8	DCM Mental	…	…	-	…	Face Validation Statistical Test		
9	DCM Physical	…	…	-	…	Face Validation Graphical Comparison		

Create plan

The plan is made by combining the result from the steps "Determine process" and "Determine techniques". It is already presented in table 4.7. The last two columns are created for the user to check whether the module/model pass the test, and make comments on verification when executing the plan. With all process and techniques to verified the cyclist simulation model determined and indicated in the plan, the next phase of the framework , "Execution of Plan", can be applied . The results of applying the according steps will be discussed in the next sub-section.

4.6 APPLICATION STEP 4: PERFORM TESTS

This subsection is dedicated to the application of step 4 in the framework. The process is to perform the techniques one by one in the order indicated in the plan. Unlike the previous steps, the user will go to some of the steps to find and fix the problem only when the user fails to complete a test. Therefore, the structure in the following paragraph will be in the order as the plan made.

Module Testing

For all the seven modules, consistency check along with face validation on the two computerised conceptual models was performed to check the correctness of the code, the logic, and the outputs according to the certain inputs. The results show that all the seven modules passed the tests, which means in the module level, the simulation model was implemented correctly without error.

Model Testing

In model testing, there are two results to be verified. One is the probability distributions of the cyclist stop position accumulated from DCM Mental, and the other is the trajectory data of the cyclists generated according to the probability distributions of steering and pedalling from DCM Physical.

One thing to be mentioned is that, for both DCM Mental and DCM Physical, there are two versions of mechanism in the simulation model to decide the final stop position (in mental layer) and final change in steering and pedalling (in physical layer) of a cyclist. One version is stochastic in the deciding process while the other is deterministic. Both of the mechanisms are based on the probability estimated by the discrete choice models. These mechanisms do not have an effect on the intermediate outputs of the modules, so they were not mentioned in the module testing. These two versions of mechanism were both implemented in the simulation model in the beginning and which one should be used in the final version for both module remains to be determined in model verification. Therefore, both the correctness of the implementation tested and the better version of the mechanism is determined in the model testing.

DCM Mental

First, a face validation on the graphical display is performed. The figure 4.1 shows the visualised results from the conceptual model (top), from the simulation model with a deterministic version of mechanism (middle), and the one with stochastic version (bottom). Each of the result shows two probability distributions, one with all cyclists (the upper one) and one with the first coming cyclist only (the lower one). The two horizontal red lines indicates the left (the upper one) and right (the lower one) boundaries of the cycle path. The vertical red line represent the stop line in the intersection, with the while dot indicating the request-green button. Each diamond implies a stop position for a cyclist, and the colour of each diamond shows the probability to be chosen with the certain colour.

From the figure, it can be observed that if the model is deterministic, only the cells with the highest probability will be chosen, which shows a big difference with the conceptual model. On the other hand, in the stochastic version, it can be observed that the result is quite similar to the conceptual model. Although the result from the stochastic version is not identical to the result from the conceptual model, we can still tell that they shared similar characteristics. For example, the most chosen position is next to the right bottom corner, where the request green button is located. Also, the stop positions in the former half (right side) of the cycle path is more often chosen than the latter half (left side). Therefore, in terms of face validation, it can be concluded that the stochastic version performs better than the deterministic one and it is adequately comparable to the conceptual model.

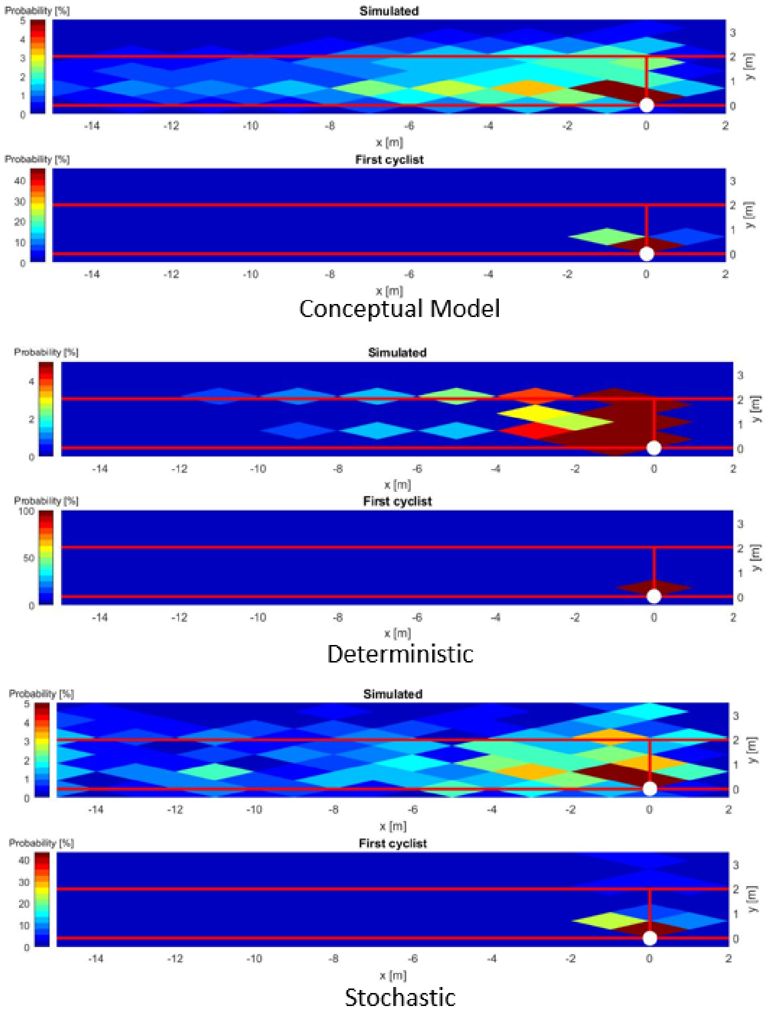

Figure 4.1: Face validation of DCM Mental using the graph to show the probability for each stop position to be chosen.

To be more certain about the result, a statistical test is performed. By applying the scheme mentioned in section 3.3.3, a specific test can be determined. The final outputs from the conceptual model is a set of probability indicating in each cell (a cell is an alternative in the choice set). A cell number can refer to a value (probability), and they are paired together. In this case, it is a comparison within the parametric test. As the main goal is to find out the differences among group means with categorical predictor variable (cell number) and the quantitative outcome variable (probability), the paired t-test is selected. To perform the statistical test, the following hypothesis is made:

H_0: The mean difference of the two samples is zero

H_1: The mean difference of the two samples is not zero

Comparing the result from the conceptual model with the result from the stochastic model, the paired t-test shows that the p-value of the test is 0.98. It is a statistically significant value ($p > 0.05$), meaning that the null hypothesis (H_0) cannot be rejected. Likewise, comparing the result from the conceptual model with the result from the deterministic model, the paired t-test shows that the p-value of the test is 0.27. It is also a statistically significant value that the null hypothesis cannot be rejected. From the results of the two statistical tests, we can conclude that both versions of the simulation model can represent the conceptual with statistical significance. However, although there is a gap between the two p-values, we cannot tell which one performs better by comparing the p-values since the larger p-value only implies less strength of evidence to reject the null hypothesis [Royall, 1986].

To conclude, the module passes the statistical tests in model testing no matter which mechanism to decide the final stop position is used. In terms of the two mechanisms, the model with stochastic version of mechanism is able to reproduce a more comparable result with the conceptual model in face validation. Also, the fact that the deterministic model generating identical results in different runs of simulation is unrealistic and it could make the use of simulation less meaningful. Therefore, the model with the stochastic mechanism performs better.

DCM Physical

Similar to DCM Mental, for DCM Physical, a face validation on the graph generated by the tool used in the conceptual model validation is performed. The tool is applied to generate the result of how often each alternative (a combination of a change in steering angle and a change in speed) is chosen and it is assumed that the simulation model should reproduce results similar to the conceptual model with the use of the tool. Figure 4.2 shows the visualised results of the probability distribution of the change in steering angle and speed for all the cyclists from the conceptual model (left), from the simulation model with a deterministic version of mechanism (middle), and the one with stochastic version (right). It can be observed that the model with deterministic mechanism shows a high concentration on the cells with the highest probability, which is quite different from the conceptual model.

On the other hand, in the stochastic version, the result is more similar to the conceptual model than the deterministic one. However, the stochastic one still shows quite different features from the conceptual model. In the circular sector, the left and right indicate the choices with negative (left) or positive (right) change in speed, while the up and down indicate the choice with change in steering angle, where up means turning left, and down means turning right. The most dissimilar part is at the right-middle edge of the circular sector, where the conceptual model shows a low probability for a cyclist to choose while simulation model shows high probability. In terms of face validation, even though the stochastic version performs better than the deterministic one, DCM physical could not be deemed as verified in model testing, therefore fail in the test.

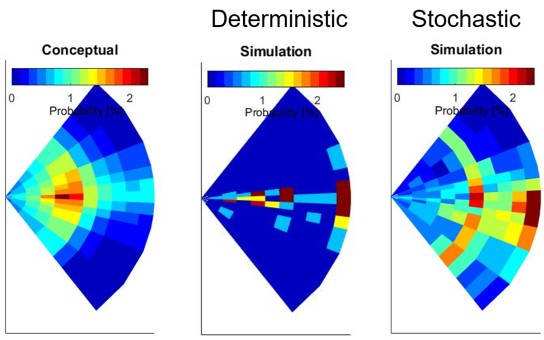

Figure 4.2: Face validation of DCM Physical using the graph to show the probability for each change in steering and pedalling to be chosen.

4.7 APPLICATION STEP 5: ANALYSE PROBLEM

As described in the framework, when the module/model fails to pass the selected test, the user needs to analyse the problem and recognise whether the problem belongs to the simulation model. The question to analyse the problem is answered.

Q3-12 Why does the module/model fail in the selected test?

Ans: The running of DCM Physical is an iterative process that every choice (change in speed and steering angle) will affect the next choice for a cyclist. Because of that, the randomness of choice will increase the uncertainty of the next choice. However, in the conceptual model validation, the result was generated independently from each position of the trajectories from the field data, so every choice that a cyclist made is independent to the choice he/she previously made. To conclude, the visualised result generated from conceptual model and simulation model are based on different assumptions, and therefore they should not be compared together.

Based on the answer to the question, it is also able to answer the next question in this sub step:

Q3-13 Is the problem belongs to the plan created in the previous phase?

Ans: Yes. The problem is caused during the selection of the test due to the lack of understanding in the existing tool and the assumptions made in the conceptual model.

As the answer to the question is ”Yes”, we can follow the framework to the next subject, ”Review plan”. Based on the understanding of the problem, the plan needs to be modified in the part of model testing. To be specific, the techniques for testing DCM Physical in model testing needs to be reconsidered.

Following the framework, the next step is to modify the plan.

4.8 APPLICATION STEP 6-2: MODIFY PLAN

This step is to fix the problem by modifying plan. The application of this step is discussed as follows.

From the problem described above, it is concluded that the face validation using the existing tool is no longer a viable test for the model testing of DCM Physical, and thus deleted from the plan. With the technique for DCM physical updated and the rest of the plan unchanged, a modified plan is presented in the table 4.8.

Table 4.8: Updated Plan for level 1 of verification

Level 1								
Order	**Module**	**Module type**	**Parameters**	**Intermediate Outputs**	**Final Outputs**	**Techniques**	**Test Passed (Y/N)**	**Other**
Module Testing								
1	Mental Choice Set Construction	Mapping module	…	…	-	Consistency Check		
2	Physical Choice Set Construction	Mapping module	…	…	-	Consistency Check		
3	Simulation Parameter Defining	Simulation module	…	…	-	Consistency Check		
4	Simulation Running	Simulation module	…	…	-	Consistency Check		
5	DCM Mental	Computerized conceptual model	…	…	-	Consistency Check Face Validation		
6	DCM Physical	Computerized conceptual model	…	…	-	Consistency Check Face Validation		
7	Visualisation	Mapping module	…	…	-	Consistency Check		
Model Testing								
8	DCM Mental	…	…	-	…	Face Validation Statistical Test		
9	DCM Physical	…	…	-	…	~~**Face Validation**~~ **Graphical Comparison**		

As the problem is solved, we can move back to "Perform tests" with the modified plan as the framework indicates.

4.9 RESUME TO STEP 4: PERFORM TESTS

After the problem is solved, it is possible to come back to this step and continue to complete the rest of the modified plan.

Graphical comparison

The DCM Physical is meant to generate the trajectories of cyclists from the intermediate results of the model. So, to make a graphical comparison, a scenario is made to generate the trajectories for testing. From the analysis of the field data (figure 4.3), it can be observed that there are an obvious pattern for the first three coming cyclists. The probability of the forth one depends more on the previous cyclists. After the forth cyclist, the pattern become less interpretable. In the data analysis, it is shown that nearly 90% of the first cyclists selected the right-front position, and 67% of the second cyclists selected the left front, queuing in the sub-line (left). With the first two cyclists selecting the front-right and front-left positions, there are of over 50% of the third cyclists selected to queue in the main (right) line, and over 60% of the fourth cyclist selected to queue in the main line also. The stop positions for the cyclists coming after are distributed widely with the probability of each position of less than 20%. Therefore, the aforementioned scenario with the first forth cyclists is chosen as it is most observed from the data. Also, it includes the important features of the queue formation model: the selection of the request-green button position for the first cyclist, the formation of the queue in the main line, the formation of the queue in the sub-line and also the interaction within multiple cyclists.

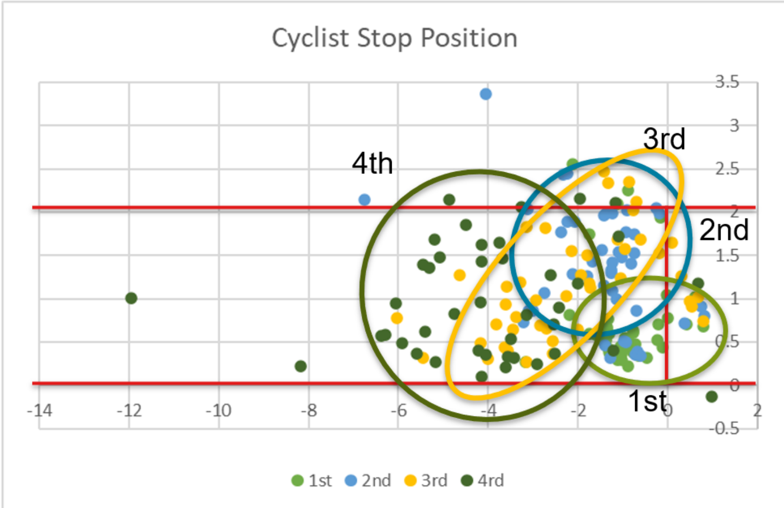

Figure 4.3: Stop position distribution for the first four cyclists

The selected scenario is presented in figure 4.4. The traffic light on the right-bottom corner indicates the request-green button. The two dash lines are used to separate the first place of each line and also identify the main line and the sub-line. The bike seats with a wheel for each and handlebars are used represent the cyclists. By observing the data, there are 7 data sets fits the scenario out of 46 sets, and a random data set is selected as a reference to make a comparison.

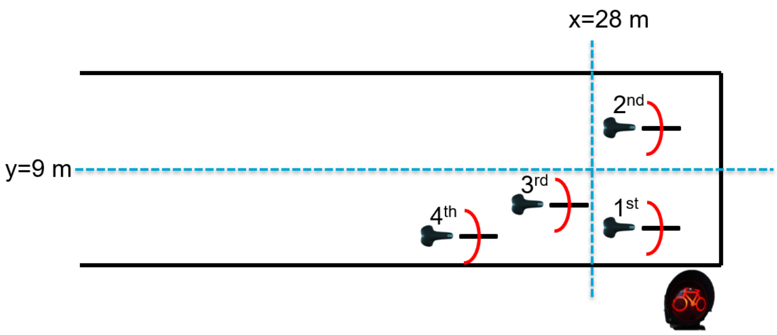

Figure 4.4: Selected scenario

The results to be compared are generated from the simulation model using the same parameters observed in the data set, which are the same initial speeds, the same initial positions, and the same aimed stop positions with similar incoming intervals.

As one run of simulation could only represent a certain case of the simulation model with the stochastic mechanism to decide the final stop position or the final steering and pedalling, a result with multiple runs, such as an average trajectory, is preferred to include different possibility that could happen. From a few trials of simulation, it is known that the average trajectory becomes a smooth line that does not resemble any of the generated trajectories over ten runs of simulation. The average trajectory generated from around five runs of simulation remains the similarity to the individual trajectories, at the same time represents the average positions from different runs of simulation. Therefore, the author decided to generate results from five runs of simulation and both the individual trajectories and the average trajectory are present together in one graph to be observed together.

The visualised results are presented in the following figures with the data observed in the system, the data from each simulation run and the average of the simulation.

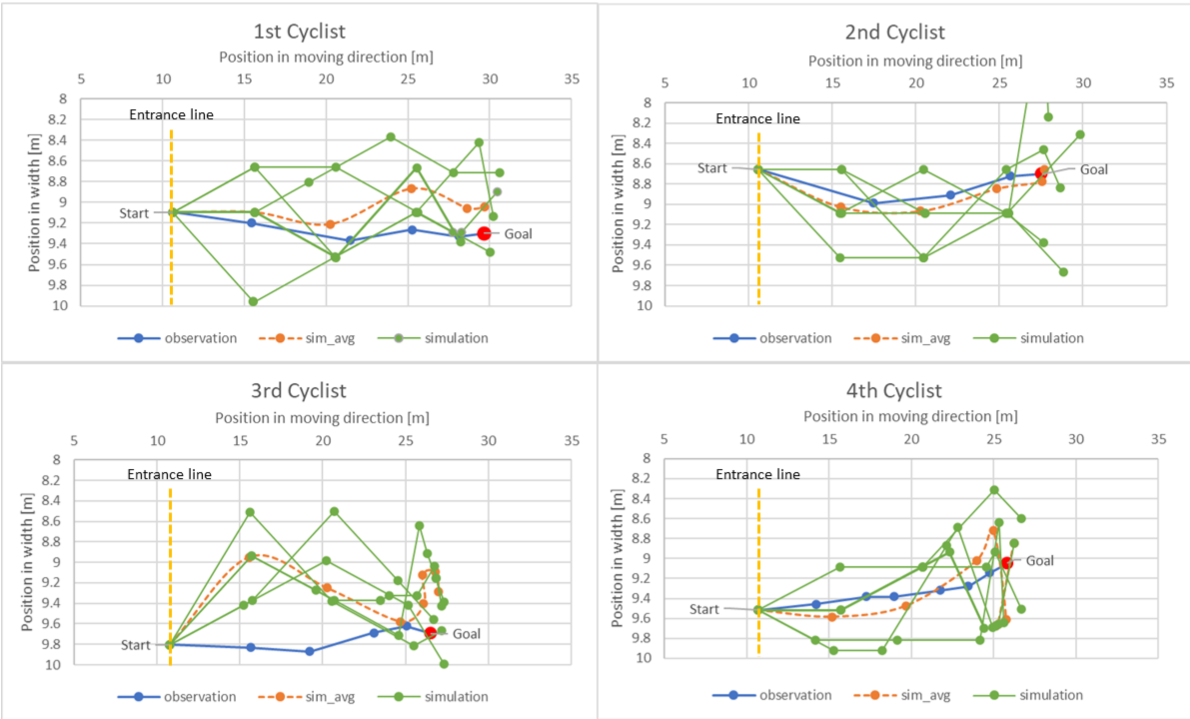

Figure 4.5: Trajectories for the first four cyclists, including the one from observation (blue), the average from five runs of simulation (dash orange), and the five individual trajectories from five runs of simulation (green).

From the average trajectories (dash lines), it can be observed that the simulated trajectories are comparable to the observed trajectories for most of the cases, except the 3rd cyclist. However, by looking into the individuals, there are quite some variations of the presented trajectories, especially for the 1st and 2nd cyclists. It might be resulted from the fact that for the first two cyclists, there are no other cyclists in front, and less interaction is needed. Therefore, the probability to choose a certain change in steering angle and speed are distributed more evenly, so it is more likely to have different trajectories for the first two cyclists. Nevertheless, in most of the case for the 1st and 2nd cyclist, the trajectories tend to converge to the aimed goal at their latter steps, which demonstrated their intention to reach the goal. On the other hand, for the 3rd and 4th cyclist, they showed more similarity between different runs, which is probably because the previous cyclists limited their movements.

To test the deterministic version of the model, the comparison between the trajectories from the stochastic model, deterministic model and field data is also made. The result is presented in figure 4.6. The trajectory generated from the deterministic model is consistent through different runs of simulation, and it behaves quite different from both the real trajectory and the average trajectory from stochastic model. In the deterministic model, only one change in steering angle is observed at the last moment, which could be unrealistic for human beings.

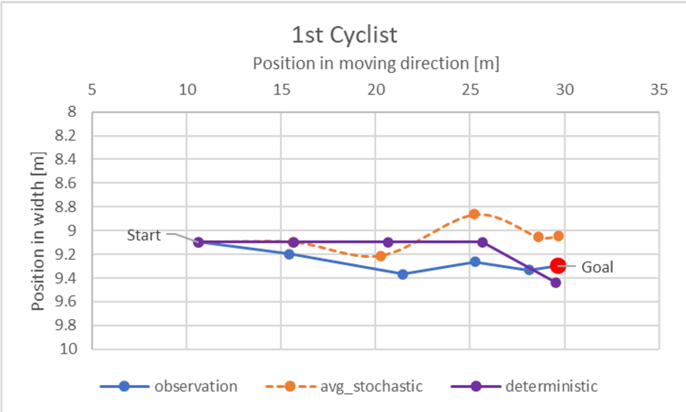

Figure 4.6: Trajectories of the 1st cyclist, including the observation (blue), the average one from the stochastic model (orange dash) and the one from the deterministic model (purple).

It is hard to tell whether the simulation model generates a valid trajectory. Both versions of the model presented the cyclists' intention to reach the goal. What could be said from the graphs is that the stochastic version of model display more comparable characteristics with the real trajectory than the deterministic one, showing some certain behaviour of the cyclists, for example the oscillation in the trajectory caused by the correction behaviour.

It is mentioned in the beginning that the verification is to compare between the conceptual model and simulation model. So the intention for this graphical comparison is not to really validate the simulated trajectories with the observed trajectories since that will be a simulation model validation. The selected observed trajectory is merely a reference for how the trajectory of a cyclist could be. The main focus of the comparison is to see whether there is a similarity between the trajectory generated from the simulation model and the real trajectory in terms of its characteristic.

From the author's perspective, what can be concluded is that the simulation model represents a part of the system, such as the cyclist correction behaviour, as the simulation is intended to do. Therefore, DCM Physical passed the model testing as it presented its intended application. However, the result is not observable in the conceptual model so it is not an objective judgement. The fact that the trajectories generated in different runs deviating with considerable distance could be a potential problem in the model.

To this point, as the plan is completed in this phase, the verification process is finished. However, during the verification of in model testing of DCM Mental, it is found that there could be a solution for the potential problem that we identified in the model. The solution is to adjust one of the parameter used in the model, and it can be tested with a local sensitivity analysis. To further develop the queue formation simulation model, there is chance to conduct a verification of level 2, where the variability of the parameters are to be tested with sensitivity analysis. The author takes it as a chance to demonstrate how the level 2 of verification could be for the user of the framework, at the same time examine the feasibility of the potential improvement to the model.

4.10 EXAMPLE OF LEVEL 2 OF VERIFICATION

When it comes to the further development of the cyclist queue formation model, the next step is to understand the applicable range of the queue formation simulation model by performing level 2 of model verification. The example of level 2 of verification is presented as follows.

In the model testing of the queue formation model, tests of face validation and statistical analysis were conducted on DCM Mental using the existing tool to generate the graphic of the probability distribution of the stop position (figure 4.1). The result of the model testing shows that the stochastic version of the simulation model prevail the deterministic version, and it is statistically significant that the simulation model can reproduce the result of the conceptual model from DCM Mental. Thus, the simulation model passed the test.

However, it can still be observed that the cyclist in the simulation model might have chosen too many cells (stop positions) with low probabilities, such as the cells on the left-hand side (at the entrance of the cycle path). Therefore, the author hypothesize that there exists a better version of the model, where the determination of the revealed stopped position is not totally stochastic. Instead, cyclists should only choose the cells with probabilities higher than a threshold. To do so, the author introduced a parameter called "level of deterministic" to the model, which will make the alternatives with little probability unavailable. With a higher level of deterministic, the cyclist will choose the cell with higher probability estimated by the discrete choice model, ignoring those cells with lower probabilities. This parameter is represented as a threshold h that the alternatives with probability lower than h are ignored. So, h = 0% means totally stochastic, the same as the current verified one. To prove that the model can be improved by introducing this new parameter, a coarse search for the best fit value is performed, as shown in figure 4.7.

The threshold h = 5% was the first guess, and it turned out to be too high. By lowering down the threshold step by step, a more comparable result was obtained with h = 1%. The similarity between the result from the conceptual model and the simulation model with h = 1% seems to be higher than with the result from the original simulation model in terms of face validation. At least the cells in the left-hand side (the entrance to the cycle path) will be less chosen with this threshold, while the other characteristic remains similar.

To explore other possibility for h, a fine search for the value of the parameter was conducted, as shown in figure 4.8. This time, the search started from h = 1% to h = 0.1%. Comparing the coarse search and fine search, we can observe that the results differ less with fewer changes in the parameter. What can be seen in the fine search is that when h is lower than 0.75%, the results become more stochastic where the cells around the edge of the cycle path become more preferred by the cyclists. Looking into the results in the fine search, it is concluded that h = 1% is the value that reproduces a more desirable result in terms of face validation.

To confirm whether the model with h = 1% really performs better than the original one, similar to what has been done in the application of the framework, a statistical test using paired t-test is performed to compared the result of the simulation model with h = 1% with the result from the conceptual model. The hypothesis for this test is the same as what was used in the application:

H_0: The mean difference of the two samples is zero

H_1: The mean difference of the two samples is not zero

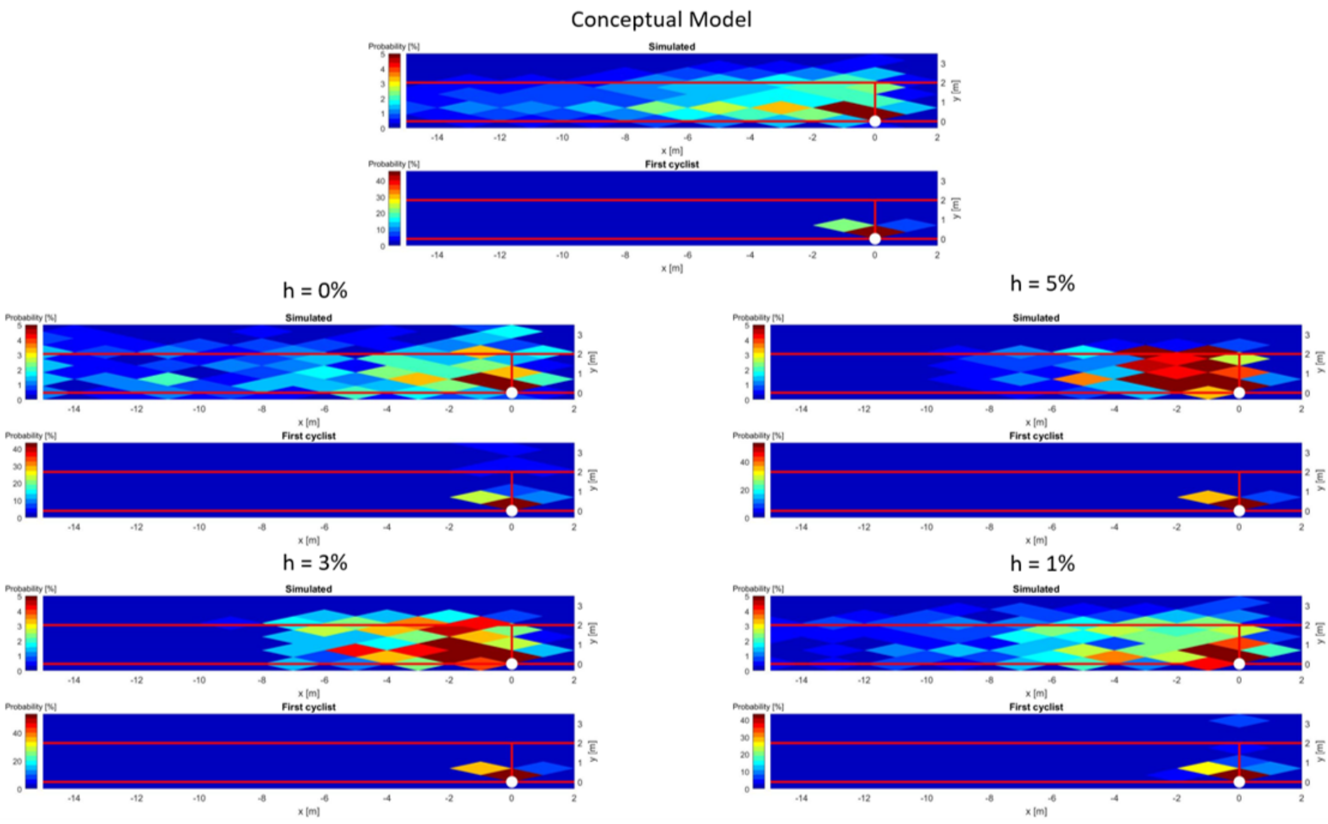

Figure 4.7: The results from DCM Mental in face validation for coarse search of a better h value.

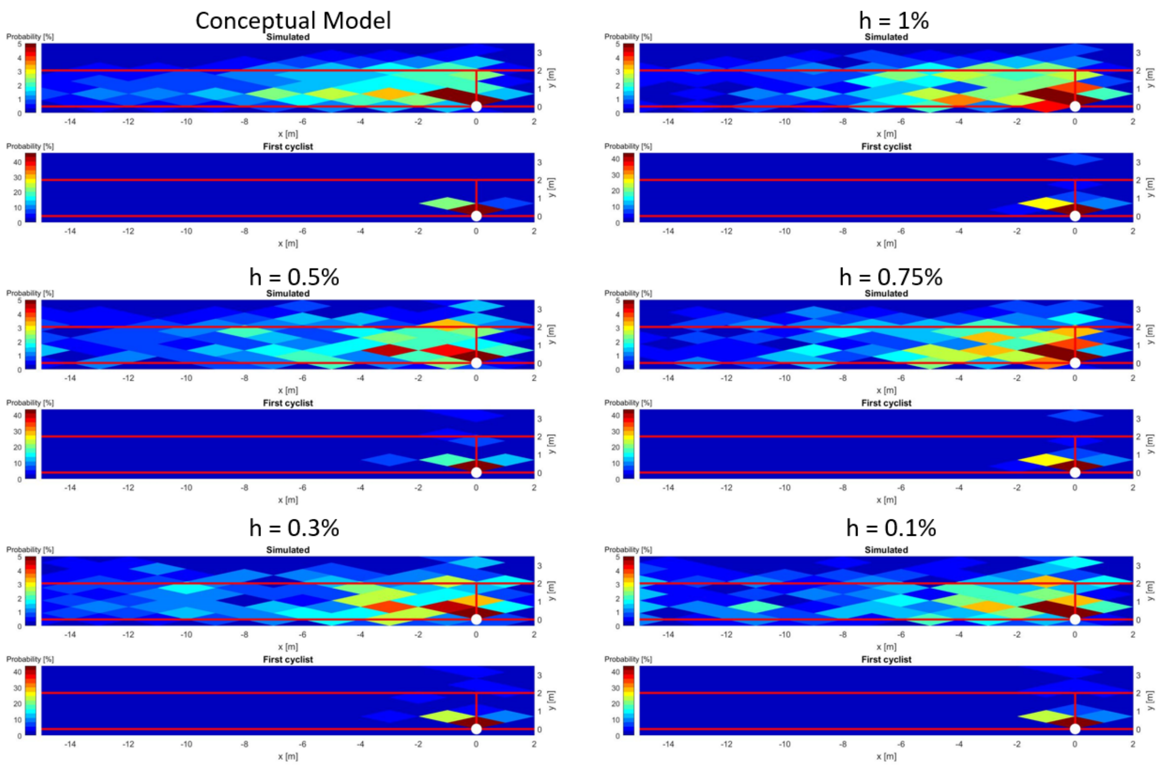

Figure 4.8: The results from DCM Mental in face validation for fine search of a better h value.

Comparing the two results with the paired t-test, the results show that the p-value of the test is 0.99. It is a statistically significant value ($p > 0.05$), meaning that the null hypothesis cannot be rejected. Sill, it does not really prove that the model with the new parameter performs better even though the p-value is higher than the one obtained from the comparison between the original simulation model and the conceptual model (0.98). What can be told is that there is more evidence to reject H_0 for the original model than the one with the new parameter [Royall, 1986].

To conclude the potential solution shows the possibility to improve the model in terms of face validation. There is a chance that the model with h = 1% performs better than the original one. However, it cannot be proved statistically.

In the example, the purpose of the verification is to examine the potential solution with the new parameter h. It turned out that the best fit value for h is 1% in terms of face validation, but it cannot be proved that the model with h = 1% performs better.

The verification of level 2 could also be practised in other ways, depending on the purpose of model verification. For example, it could be finding the workable range of h, or it could be an optimisation of the model by tuning h, depending on the requirement for the level 2 of verification.

4.11 END OF FRAMEWORK APPLICATION

As the plan is completely executed and an example to perform level 2 of verification is demonstrated, this part of the research, ”Framework Application”, is finished. In the next chapter, the assessment of the framework based on the results of the application are to be discussed.

5 ASSESSMENT OF THE FRAMEWORK APPLICATION

In this chapter, the assessment of the framework is discussed based on the results of the framework application to the queue formation simulation model. The pros and cons of the verification framework and the suggestion to the tested simulation model are also deliberated.

First, the assessment based on whether the framework is successfully applied to the queue formation simulation model is addressed. In that part, the discussion on whether it is possible to fully perform the framework and the evaluation of the product from the framework are discussed. In second part, the pros and cons of the framework is elaborated based on the experience learned from the framework application. The pros and cons of the framework found in the application are listed and explained. Last, the room for improvement of the queue formation simulation model is addressed.

5.1 ASSESSMENT OF THE FRAMEWORK

To assess the framework, two indicators that Leutscher [2018] used to assess a new methodology are applied: 1) the ability to perform all steps of the framework and 2) qualification of the product from the framework. In the following paragraph, the assessment based on these two criteria is discussed.

5.1.1 Completeness of application

In terms of the ability to perform all the steps of the framework, the framework application is summarised in table 5.1 by using the check points at each step/phase in the framework. In this table, we can see that most of the steps and the sub-steps were performed. The only step and sub-step that were not performed are "Modify simulation model" along with one of its sub-step "Restart framework". The reason that the step and the sub-step were not performed is that they are not necessary in this case. Also, an example to perform a small part of testing for level 2 of verification in phase 2 was demonstrated successfully. Therefore, from the experience gained from the verification of the queue formation simulation model, it is known that all of the required steps can be performed without too much effort. Still, the framework is mainly performed for a verification of level 1, and whether the framework provides adequate information for a verification of level 2 is unclear. To conclude, based on the experience learned from the framework application, the framework has proved its ability in application with all the required steps performed.

5.1.2 Product of the framework

To assess the product of the framework, the characteristics that a verified simulation model should include are used. As mentioned in the literature review, there are two objectives of model verification: 1) ensure the basic functionality of the simulation model; and 2) decide the applicable range of the simulation model. In terms of the

Table 5.1: Framework steps performed

Framework steps	Performed?	Other comment
Start	Yes	
Phase 1: Creation of Plan		
Step 1: Determine level of verification	Yes	
Recognise purpose	Yes	Derived from the study of the conceptual model
Determine level	Yes	Level 1 is selected according to the purpos
Step 2: Module Identification	Yes	It really requires the source code or the specificatic
Recognise module	Yes	It takes time. Could be iterative.
Specify effect	Yes	Almost done together with the first subject.
Classify module type	Yes	The classification is useful.
Identify parameters	Yes	Hard to distinguish the relevent and irrelevent ones
Identify outputs	Yes	
Indicate relationship	Yes	Quite intuitive
Step 3: Verification Planning	Yes	Rather fast when step 2 done properly
Determine process	Yes	
Determine techniques	Yes	The techniques could be differents for different use
Create Plan	Yes	
Phase 2: Execution of Plan		
Step 4: Perform tests	Yes	DCM Physical failed in face validation of model testing.
Continue plan	Yes	
Step 5: Analyse problem	Yes	DCM Physical: possible reason concluded.
Step 6-1:Modify simulation model	No	No modification needed due to the fact the the problem is resulted from the planning, not the model.
Restart framework	No	Not required in this case.
Step 6-2: Modify plan	Yes	Change made in techniques for DCM Physical.
Review plan	Yes	Change in DCM Physical required.

first objective, there are two parts to be discussed. The first part is the basic function of each part of the model, which implies the correctness of the implementation of the simulation model. The second part is the ability of the simulation model to reproduce the result of the conceptual model or the ability of the simulation model to perform its intended application as a whole. For the second objective, it is about whether the applicable range of the simulation model is identified, including the boundaries and the suggested values of the parameters used in a model. By applying these criteria, the product from the framework can be assessed.

The aforementioned criteria are used in table 5.2 to assess the queue formation simulation model after testing with the framework. The "v" in the table means that the module/model has proved that it met the certain criteria during the testing. From the table, we can observe that the basic function of each module of the queue formation simulation model is guaranteed through module testing during the framework application. Furthermore, in the model testing, DCM Mental displayed its ability to reproduce the result of the conceptual model and DCM Physical demonstrated the cyclists' movement behaviour with a comparable result to the real trajectory. Based on these two points, the tested queue formation simulation model from the framework application has reached the standard of the first objective of model verification, thus could be said as a verified simulation model within the study area where the conceptual model is proposed.

Table 5.2: Product assessment

Module	Basic function	
Module Level		
Mental Choice Set Construction	v	
Physical Choice Set Construction	v	
Simulation Parameter Defining	v	
Simulation Running	v	
DCM Mental	v	
DCM Physical	v	
Visualisation	v	
Model Level	**Reproduce result?/ Perform as intended?**	**Applicable range recognised?**
DCM Mental	v	
DCM Physical	v	

5.2 PROS AND CONS OF THE FRAMEWORK

During the framework application, the pros and cons of the framework are recognised, and they are summarised and explained in this section.

5.2.1 Advantages

The main advantage of the framework is that it provides a systematic way to verify a traffic simulation model, which results in the following benefits to use this framework:

- The framework is well organised, and it is easy to follow and therefore there is a chance to reduce time spending on model verification
- The framework is generic for different purposes of model verification. Both purposes of verification found in the literature are considered by the framework.
- The verification process can be traceable. The steps to make a plan, and the steps to execute the plan are explicitly indicated in the framework, so whenever a mistake is made, the user can trace back to find out the problem. Also, for other users, the verification of a certain model can be studied as documentation is emphasised in the framework.
- The verification process and the product become more comparable. As the two levels of model verification are separately discussed, it is possible to compare

the verification process between different models according to the level of verification. Verified simulation models are also easier to be compared with the models fulfilled the same level of verification.

5.2.2 Disadvantages

From the experience of the application, there was no structural disadvantage found in the framework. However, there are some limitations that observed during the application and they are summarised as follows:

- The framework is limited to the conceptual model. Although the necessity of the conceptual model to conduct verification is explained, the author believes there could be a model that requires verification without access to the conceptual model. In that case, sensitivity analysis might still be useful in verification, but it is not included in the framework.
- The framework is still limited to the type (of traffic modes) of simulation model. The framework is built based on the knowledge of verification of vehicular simulation model, of which the most methodology applications are found in the literature. It is created for traffic simulation model with the discussion on the difference to verify a cyclist simulation model. However, its applicability to other traffic simulation model is not really discussed.
- The step "Module Identification" could be challenging for some of the models. In the framework, this step could performed by looking into the specification or the source code of the simulation model or by categorising the required input when the model specification or the source code is not available. It is found that this step is easier to be performed with the specification or the source code available during the application. However, the source code and the specification are not accessible for most of the simulation model, as it could be the intellectual property for some companies or developers. Even the source code is available, a user could spend excessive time to understand the code, which could lead to increase in cost for model verification.

5.3 SUGGESTION TO THE QUEUE FORMATION SIMULATION MODEL

Besides the qualification of the input simulation model, another objective of verification is to identify the potential problem in the tested simulation model for the model user/developer to understand the problem and further improve the model.

5.3.1 Room for improvement in DCM Physical

The module "DCM Physical" is a computerised conceptual model in the queue formation simulation model specified from the physical layer discrete choice model of the queue formation conceptual model. In the conceptual model, the physical layer discrete choice model is used to describe the "choice preference" of each cyclist to a certain change in speed and steering angle. Its application in the simulation model is to generate a trajectory for each cyclist according to those preference.

During the model testing of the simulation model in the framework application, it is found that DCM Physical is able to describe the behavior of the cyclists, such as the correction movement and the intention to reach the goal. The generated average trajectories were similar to the observed trajectories. However, the trajectory of each cyclist generated from each run of the simulation would diverge from the average trajectory with a considerable distance in some runs. It is hard to say that the deviation is within a reasonable range. Even though the model passed the tests of verification, it might need some improvement for the further validation of the model.

A possible reason for this problem is the long update interval in the model. The time step used currently in the simulation model is 1 second, the same as the conceptual model, which could result in a long traveled distance for a cyclist before a new decision is made. Therefore, one of the potential solution is to use a smaller time step for updating the model so that the simulated cyclists could perform their correction behavior more instinctively as a real cyclist. However, how the model will perform when changing the value of the time step remains a question, and it is to be discovered with a verification of level 2.

To conclude, in this section, the potential problem of the simulation model is indicated. The problem is only found in module "DCM physical", and the rest of the model is performing well from the observation of the framework application.

5.4 CONCLUSION

In this chapter, the framework is assessed with the results from the framework application. The pros and cons of the framework, and the room for improvement of the simulation model are also discussed. Based on these, it is possible to answer the according sub-questions:

Q1-5 How is the framework performed in the application?

Ans: From the results of the framework application, it is known that all of the required steps in the framework can be performed. The product of the framework, in this case the queue formation simulation model, is also qualified as a verified simulation model (within the study area where the conceptual model is proposed). In conclusion, the framework demonstrated its intended practicability in the application.

Q1-6 What are the pros and cons of the framework that discovered in the application?

Ans: For the pros, the framework is systematically structured which results in a well organised and traceable verification process with its procedure and product comparable. In terms of the cons, the framework is limited to the existence of conceptual model, the type (of traffic mode) of the input model, the availability of the source code or model specification (of the input simulation model), and the ability of the user to under stand the source code (if available).

Q1-7 If there is room for improvement of the simulation model, what is it?

Ans: There is one potential problem found during the framework application, which is the considerable deviation of the generated trajectories from each run of the simulation. This deviation could be caused from the relative large time step to update the model and using a smaller number could be a solution.

To conclude, the application of the framework provides adequate information to answer the sub-questions related to the assessment of the framework and the product. Based on these answers, we can draw the conclusion of this research, which is discussed in the next chapter.

6 CONCLUSIONS, DISCUSSIONS AND RECOMMENDATIONS

This chapter is dedicated to the conclusions, discussions and recommendations of this research. First, the conclusions of this research are addressed as the implications for science and practice. After that, the discussion on this research and the recommendations for future research, will be presented.

6.1 CONCLUSIONS

The goal of this research is to create a framework that can be used as a standard procedure to verify different traffic simulation models, including cyclist simulation models. This goal is done by exploring the existing methodologies that used in the verification of traffic simulation model and deriving the essential steps from them in an organised structure with a discussion on the possible difference to verify a cyclist simulation model. To assess the framework, it is applied to the cyclist queue formation simulation model and the results of the application are analysed.

Based on the research objectives, the main research question and the seven sub-questions have been formulated. As the answers to all the sub-questions are explored in the according parts of the research, it is possible to answer the main research question. In this section, the research findings, covering the answers to the sub-questions, are discussed first. Following that, the answer to the main research question is deliberated.

6.1.1 Research findings

With a thorough exploration of the existing literature, the definition of model verification is defined as a process to justify the implementation of the simulation model by comparing the simulation model with the conceptual model using system data. Two purposes of model verification are categorised: 1) ensure the basic functionality of the simulation model; and 2) decide the applicable range of the simulation model.

The understanding of the definition and the purpose of model verification helps to identify the methodologies used in the verification of traffic simulation model. There are three parts of methodologies that identified in this research: the process, the techniques, and the decision on the process and techniques.

In terms of process, there are two types of process found in the literature: 1) a series of tests assessing the functionality of the simulation model by comparing the simulation model with the conceptual model; 2) a series of sensitivity tests on the crucial parameters, assessing the influence of the parameters on the model performance. In terms of techniques, there are: animation, consistency check, face validation, functional testing, graphical comparison, sensitivity analysis, special input testing statistical test, and tracing found used in the literature. In terms of decision for process and techniques, it is known that the process is decided by the purpose of verification, and the techniques are determined the purpose of verification, decision approach, and system observable (possibility to collect the required data).

Besides, the potential difference to verify a cyclist simulation model is briefly addressed based according to the difference between the microscopic cyclist simulation models and microscopic vehicular simulation models. It is known that the possible difference could come from the different dimension of the modeling behavior, which makes the justification of the output from vehicular simulation model more intuitive than cyclist simulation model. However, the difference does not effect much on the verification process.

According to the methodologies found in literature and the difference to be considered for the cyclist simulation models, the core actions to be included in the framework are developed. As discovered in the literature, there are two phase in verification: "Creation of Plan" and "Execution of Plan". The steps to be included in phase "Creation of Plan" are 1) "Determine level of verification", 2) "Module Identification" and 3) "Verification Planning". And the steps to be included in phase "Execution of Plan" are 4) Perform tests, 5) Analyse problem, 6-1) Modify simulation model, and 6-2) Modify Plan.

The developed framework is assessed through an application to the cyclist queue formation simulation model. With the application, the framework displayed its practicability with the ability to perform all the required steps and the qualification of the product, a verified simulation model. In the application, the pros and cons are also discussed. The framework provide a systematic measure for model verification that produces a traceable process and comparable product (verified simulation model). It can be applied for different purposes of verification, and no structural disadvantage is found. The only limitation of the framework is the requirement of the conceptual model.

Regarding the verified cyclist queue formation simulation model, there are a few points to be addressed. First, the simulation model proved its credibility in verification of level 1, passing all the required tests. In module testing, the correctness of the implementation and consistency of the intermediate results are checked. For model testing, DCM Mental proved its ability to reproduce the results of the conceptual model with a statistical test, and DCM Physical also demonstrated its intended application when exploring its model behaviour. In addition, the stochastic mechanisms to decide the final choice for both DCM Mental and DCM Physical surpassed the deterministic one, producing a more comparable result to the conceptual model and the intended behavior.

Second, the verification of DCM physical was a challenge. According to the study of the conceptual model [Gavriilidou et al., 2019b], the result from the discrete choice model in "Physical Layer" in the conceptual model (which could be referred to the module "DCM Physical" in the simulation model) is generated based on the assumption that each choice of the change in steering and pedalling of a cyclist is independent from the previous choice. However, in the simulation model, a trajectory is generated by a set of continuous choices that each of them is made according the situation caused by the previous choice. Therefore, the result of the module "DCM Physical" should not be compared the result from the conceptual model.

Because of that, instead of comparing to the conceptual model, the trajectories generated by the simulation model is tested with a comparison to the real trajectories as the intended outputs. It is found that the stochastic model display more comparable characteristics with the real trajectory than the deterministic one. Also, some certain behaviour of cyclists are observed from the simulation, such as the oscillation in the trajectory caused by the correction behaviour and the cyclists' intention to reach the goal. From the observation of DCM Physical, it is concluded that the simulation model represents a part of the system. However, without an objective judgement, such as a statistical test, the validity of the generated trajectories remains a question.

Last, there is a potential problem in the model found during the verification. The trajectories generated from the DCM Physical would deviate with considerable distance in different runs of simulation. This could caused difficulty in the further development of the model, such as model validation. A potential solution to the problem is to lower down the update interval used in the model. However, this solution requires the understanding of the effect caused by each parameter, which necessitates a verification of level 2 for the queue formation simulation model.

6.1.2 Answers to the research question

The goal of this research is to find an answer to the main research question:

What is the generic framework to verify traffic simulation models with the consideration of cyclists that hitherto has not been discussed?

Based on the aforementioned research findings, it is possible to answer the main research question. The framework proposed in this research (figure 0.1) is built to be the answer to the research question. Through the research, with a concrete development of the framework according to the existing methodologies and a comprehensive assessment based on the application, the proposed framework has proved its feasibility to be the answer to this question. Therefore, the framework is the generic framework to verify traffic simulation models with the consideration of cyclists.

6.2 DISCUSSIONS

In this research, a framework is created based on the methodologies used to verify traffic simulation models. It is created to represent a standard procedure to verify traffic simulation model with consideration of cyclists. This framework has been assessed through the application to the queue formation simulation model developed based on the queue formation (conceptual) model of [Gavriilidou et al., 2019b].

During the application of the framework, it was found that it is possible to apply the framework to verify the cyclist queue formation simulation model, currently being developed by the Active Mode Lab in TU Delft, with all steps performed and the product reaching the criteria. Based on this experience, it can be concluded that a repeat of this research could result in similar results. However, the process and techniques may differ due to the different preference of the framework users.

The framework is built on a number of the previous studies of traffic simulation model verification, and it explains the performance of the framework. The proposed framework and its applicability to verify a cyclist simulation model are assumed to be complementary to the missing study on the standard procedure of model verification, where cyclist simulation models could also apply. Therefore, it is suggested that this framework should be considered when verification of any traffic simulation model is required.

However, it should be discussed that this framework is only applied to a single case of the queue formation model, so the current framework is only assessed with the characteristics of this case. A larger amount of cases where the framework is applied could make it possible to include more characteristics of other simulation models and improve each step and sub-step of the process in different aspects. By performing a comparative case study, it would be able to find and explain similarities and differences of the framework application. Therefore more research with this framework is recommended to determine the feasibility of the framework to verify other traffic simulation models.

6.3 RECOMMENDATIONS

Lastly, in this section, the recommendations for future research are made. There are four recommendations made for the future direction. The first and second ones are in the aspect of further development of the verification framework. The third one is about the validation of the framework with more case studies. The fourth one discusses the completion of the cyclist queue formation simulation model. The recommendations are addressed below:

i The first recommendation for future research is that the planning for the different levels of verification could be extended. As described in the framework, the plannings for level 1 and level 2 of verification are different, especially when selecting the techniques. The focus of this research is on the structure of the framework, so neither the selection of the statistical analysis for level 1 of verification nor the design of sensitivity analysis for level 2 of verification, was comprehensively exploited during the development of the framework. However, both of them are important to the completeness of the framework. Including any of the topics to the current framework and discussing the extra needs of the framework due to the change will be credible for the development of the framework, and also for the user and researcher of the framework.

ii Another recommendation would be the specification for different types of simulation model. In this research, the author tried to identify the difference between verification of cyclist simulation model and verification of vehicular simulation model. However, the conclusion made in the literature review did not really become a part of the framework explicitly. As discussed in the literature review, the possibility for the vehicular simulation model to conduct verification by comparing other existing models, as a reference to an intended (realistic) behaviour of the modelling system, could be helpful. If the characteristics of different types of simulation model, such as the dimension of the indicators we discussed, are further explored and categorised, there will be a higher chance to justify a model with other models. Thus, the limitation requiring the conceptual model of the framework could be released.

iii The third recommendation is to conduct validation of the framework by applying the framework to other traffic simulation models. In the discussion, it is mentioned that the new framework has only applied to one case, the queue formation simulation model. Although it is concluded that the framework can be applied to the cyclist simulation model and verify the model in this case, there is still a chance that the framework fails to perform verification for other simulation models. More cases where the framework is applied are needed as a way to validate the framework at the same time provide improvements to the framework. The comparison between and within the verification process and product of different types of simulation models, in terms of mode of transport or scale of the model (micro, macro, meso), could also be helpful to complete such a general method for model verification.

iiii When it comes to the further development of the cyclist queue formation model, the next step is to understand the applicable range of the queue formation simulation model by performing a model verification of level 2. The necessity for the verification of level 2 is brought by the potential solution to the identified potential problem in the model, which is to lower down the update interval (time step ”dt”). This potential improvement for the DCM Physical can be further analysed when the influence of each parameters on the model are explored.

BIBLIOGRAPHY

S. Algers, E. Bernauer, M. Boero, L. Breheret, C. Di Taranto, M. Dougherty, K. Fox, and J. Gabard. Review of micro-simulation models. review report of the smartest project; deliverable 3, 1997, 1997.

O. Balci. Verification, Validation, and Testing. 1998.

R. F. Benekohal. *Development and validation of a car following model for simulation of traffic flow and traffic wave studies at bottlenecks*. PhD thesis, The Ohio State University, 1986.

R. R. F. Benekohal. *Procedure for validation of microscopic traffic flow simulation models*. Number 1320. 1991.

H. Botma and H. Papendrecht. Traffic operation of bicycle traffic. *Transportation Research Record*, (1320), 1991.

J. Dekoster and U. Schollaert. *Cycling: the way ahead for towns and cities*. Office for Official Publications of the European Commission, 2000.

Eco-Counter. 2019 worldwide cycling index. https://www.eco-compteur.com/en/2019-worldwide-cycling-index/, 2019.

A. Gavriilidou, W. Daamen, Y. Yuan, N. V. Nes, and S. P. Hoogendoorn. Empirical findings on infrastructure efficiency at a bicycle T-junction. pages 1–15.

A. Gavriilidou, W. Daamen, Y. Yuan, and S. P. Hoogendoorn. Modelling cyclist queue formation using a two-layer framework for operational cycling behaviour. *Transportation Research Part C: Emerging Technologies*, 105(June):468–484, 2019a. ISSN 0968090X. doi: 10.1016/j.trc.2019.06.012.

A. Gavriilidou, M. J. Wierbos, W. Daamen, Y. Yuan, V. L. Knoop, and S. P. Hoogendoorn. Large-Scale Bicycle Flow Experiment: Setup and Implementation. *Transportation Research Record*, 2673(5):709–719, 2019b. ISSN 21694052. doi: 10.1177/0361198119839974. URL https://doi.org/10.1177/0361198119839974.

E. J. Gonzales, N. Geroliminis, M. J. Cassidy, and C. F. Daganzo. On the allocation of city space to multiple transport modes. *Transportation planning and technology*, 33(8):643–656, 2010.

S. Handy, B. Van Wee, and M. Kroesen. Promoting cycling for transport: research needs and challenges. *Transport reviews*, 34(1):4–24, 2014.

D. Helbing, A. Hennecke, V. Shvetsov, and M. Treiber. Micro-and macro-simulation of freeway traffic. *Mathematical and computer modelling*, 35(5-6):517–547, 2002.

S. P. Hoogendoorn and P. H. Bovy. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 215(4):283–303, 2001.

R. Horiguchi and M. Kuwahara. Verification process and its application to network traffic simulation models. *Journal of Advanced Transportation*, 36(3):243–264, 2002. ISSN 01976729. doi: 10.1002/atr.5670360304.

L. Huang, J. Wu, F. You, Z. Lv, and H. Song. Cyclist social force model at unsignalized intersections with heterogeneous traffic. *IEEE Transactions on Industrial Informatics*, 13(2):782–792, 2017. ISSN 15513203. doi: 10.1109/TII.2016.2597744.

H. Jiang, T. Wen, P. Jiang, and H. Han. Research on cyclists microscopic behavior models at signalized intersection. In *Road safety on four continents: 16th international conference*, 2013.

F. Kessels. *Mesoscopic Models: Introduction to Traffic Flow Theory Through a Genealogy of Models*, pages 99–106. 01 2019. ISBN 978-3-319-78694-0. doi: 10.1007/978-3-319-78695-7_6.

J. P. Kleijnen. Verification and validation of simulation models. *European journal of operational research*, 82(1):145–162, 1995.

V. Komkov, K. K. Choi, and E. J. Haug. *Design sensitivity analysis of structural systems*, volume 177. Academic press, 1986.

B. Leutscher. Methodology for crowd management measures: A study to define and apply a methodology to create crowd management measures suited for hubs under (re) constructions while business is performed as usual. 2018.

X. Liang, M. Baohua, and X. Qi. Psychological-physical force model for bicycle dynamics. *Journal of Transportation Systems Engineering and Information Technology*, 12(2):91–97, 2012.

M. D. Lowe. Alternatives to the automobile: Transport for livable cities. *Ekistics*, pages 269–282, 1990.

X. Ma and D. Luo. Modeling cyclist acceleration process for bicycle traffic simulation using naturalistic data. *Transportation Research Part F: Traffic Psychology and Behaviour*, 40:130–144, 2016. ISSN 13698478. doi: 10.1016/j.trf.2016.04.009. URL http://dx.doi.org/10.1016/j.trf.2016.04.009.

J. Morio. Global and local sensitivity analysis methods for a physical system. *European journal of physics*, 32(6):1577, 2011.

B. K. Nayak and A. Hazra. How to choose the right statistical test? *Indian journal of ophthalmology*, 59(2):85, 2011.

D. L. Parnas. A technique for software module specification with examples. *Communications of the ACM*, 15(5):330–336, 1972.

D. Pilgrim. Model evaluation, testing and parameters estimation in hydrology. 1975.

H. Rakha, B. Hellinga, M. Van Aerde, and W. Perez. Systematic Verification, Validation and Calibration of Traffic Simulation Models. *Proceeding of the 75th Annual Meeting of the Transportation Research Board*, 1996.

R. M. Royall. The effect of sample size on the meaning of significance tests. *The American Statistician*, 40(4):313–315, 1986.

A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto. *Sensitivity analysis in practice: a guide to assessing scientific models*, volume 1. Wiley Online Library, 2004.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

R. G. Sargent. Verification and validation of simulation models. *Proceedings - Winter Simulation Conference*, pages 166–183, 2010. ISSN 08917736. doi: 10.1109/WSC.2010.5679166.

S. Schlesinger. Terminology for model credibility. *Simulation*, 32(3):103–104, 1979.

D. Taylor and H. Mahmassani. Bicyclist and motorist gap acceptance behavior in mixed-traffic. In *Proceeding of the 78th annual meeting of the transportation research board, Washington*, 1998.

M. A. Taylor. Evaluating the performance of a simulation model. *Transportation Research Part A: General*, 13(3):159–173, 1979.

V. Trajkovski. How to select appropriate statistical test in scientific articles. *Journal of Special Education and Rehabilitation*, 2016.

J. Vasic and H. J. Ruskin. Cellular automata simulation of traffic including cars and bicycles. *Physica A: Statistical Mechanics and its Applications*, 391(8):2720–2729, 2012.

N. Zenina and Y. Merkuryev. The basis for the transportation microscopic simulation model validation. In *2019 60th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, pages 1–6. IEEE, 2019.

A APPENDIX: FULL GRAPHS AND TABLES

Table A.1: Full parameters in the queue formation simulation model

Module	Parameters
Mental Choice Set Construction	Xmin, Xmax, Ymin, Ymax, dx, dy, Shape, Xg, Yg
Physical Choice Set Construction	max_acc, max_steer, min_steer, (dt), Fgrid, vgrid, Fg, vg
Simulation Parameter Defining	Inflow, bikeType, vxDesire, vyDesire, vMax, vxMax, vyMax, Cycletime, Tstep, (dt)
DCM Mental	Betas, Xg, Yg, Infrastructure, ID, agentBikes, Choice
DCM Physical	Betas, Vg, Fg, Infrastructure, ID, agentBikes, Choice, dt, t
Simulation Running	Newflow, InflowRem, NoNewBike, NoBike, agentBikes, Choice, Tk, Tstep, trajectories_x, trajectories_y
Visualisation	xmin, xline, yroad, ycurb, trajectory_x, trajectory_y

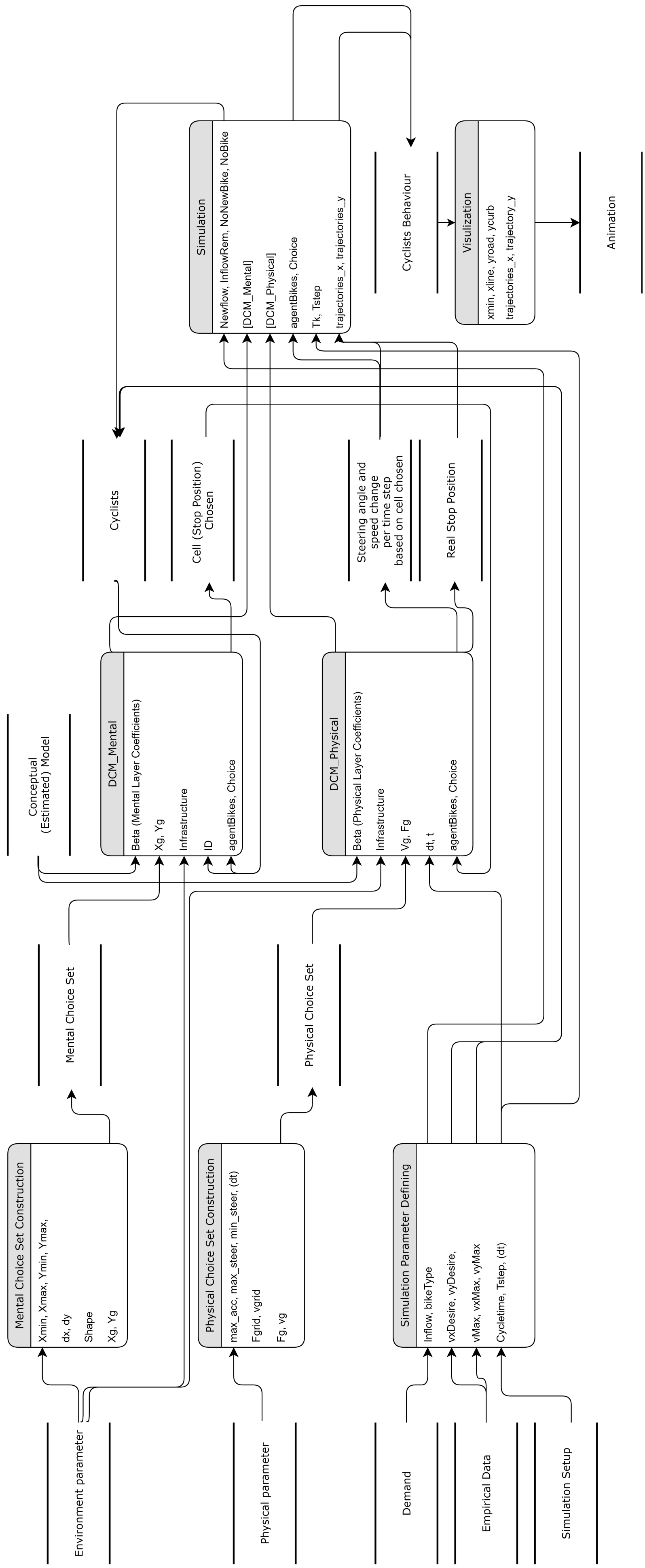

Figure A.1: Visualized structure of queue formation simulation model

Table A.2: Completed table of module identification for the queue formation simulation model

Module	Effect	Module type	Parameters	Intermediate Outputs	Final Outputs	Relationship
Mental Choice Set Construction	Defined the mental layer choice set	Mapping module	Environment Parameter (Xmin, Xmax, Ymin, Ymax)	Mental Choice Set (Xg, Yg)	-	Output to DCM Mental
Physical Choice Set Construction	Defined the physical layer choice set	Mapping module	Physical Parameter (αmax, θmax)	Physical Choice Set (Fg, Vg)	-	Output to DCM Physical
Simulation Parameter Defining	Initialisation of the parameters used for simulation	Simulation module	Simulation Setting (Timestep, cycle time, Vmax, Vdesire)	-	-	
DCM Mental	The discrete choice model that describe the preference to the stop position of a cyclist	Computerized conceptual model	Choice set, Attributes, Coefficients (Xg, Yg, agentBike, Betas)	Preference to a stop position for a cyclist	Probability distribution of the cyclists stop position	(Final result) output to Visualisation
DCM Physical	The discrete choice model that describe the preference to the change in steering angle and speed of a cyclist	Computerized conceptual model	Choice set, Attributes, Coefficients (Xg, Yg, agentBike, Betas)	Preference to a state for a cyclist	Trajectories of the cyclists	(Final result) output to Visualisation
Simulation Running	1. Run the simulation according to the simulation setting. 2. Generating cyclists 3. Handling the input and output to and from the DCM Mental and DCM Physical	Simulation module	Timestep, Inflow, Cyclist state (Tk, Newflow, agnetBike)	New cyclists (agnetBike)	-	Output to DCM Mental and DCM Physical
Visualisation	Display the animation of the results.	Mapping module	Simulation result (trajectory_x, trajectory_y)	Animation	-	

COLOPHON

This document was typeset using LaTeX. The document layout was generated using the arsclassica package by Lorenzo Pantieri, which is an adaption of the original classicthesis package from André Miede.

TU Delft