Masters of Science in Civil Engineering - Transport and Planning
MSc Thesis at Delft University of Technology

Evaluating Macroscopic DTA Models – For Who, When and How?

Aswin Menon Nandakumar

Student Number: 4987012

Thesis Committee

Committee Chair	: Prof.dr.ir. SP(Serge) Hoogendoorn
Daily Supervisor (TU Delft)	: Dr.ir. H. (Henk) Taale
Daily Supervisor (Dat.Mobility)	: Feike Brandt and Luuk Brederode
External Faculty member (TU Delft)	: Prof.dr.ir. B.H.K. (Bart) De Schutter

December 10, 2021

Cover Page Design Courtesy: Anupama Nandakumar, 2021

Preface

This thesis report marks the end of my master's degree in Civil Engineering in Transport and Planning, which I started in Sep-19. The journey was surely uplifting in all essence helping me achieve my ambition to become a transport professional.

I would like to acknowledge the contributions, time and effort taken by my honorary thesis committee. First of all, I would like to thank Prof.dr.ir. Serge Hoogendoorn, who agreed to become the chairman for my committee within his busy work schedule. Your comments and feedback throughout the course of this project made me identify aspects which were easily overlooked. The constant discussions with my daily supervisor at the University Dr. ir. Henk Taale, helped me align and shape the problem statement at hand to a valuable research topic. I would also like appreciate his quick responses and the ease with which I could approach him.

This thesis would not be possible if not for the constant support provided by my daily supervisors at Dat. Mobility Ir. Luuk Brederode and Ir. Feike Brandt. I would like to thank both of them for offering me an internship position. The ease with which I could ask questions to Luuk made me realize his genuine interest in my personal growth and in molding the thesis project to a quality work product. I would also like to appreciate his honest and critical comments which also shows his immense patience. Feike was my one-stop shop for all my practical issues with both software and hardware. The speed with which the testing kicked-off would not be possible without his quick responses and physical project room sessions at the Deventer office of Dat. Mobility. I would also like to thank the company Mint NV, Belgium for providing me with their transport network for running the scalability tests.

The non-conventional perspective offered by my external supervisor at the university Prof.dr.ir. Bart De Schutter was beneficial for me in understanding different prospects of this research. I would like to thank you for your effort and time considering your busy schedule in the 3Me department.

After my tenure as a civil engineer in the construction industry for four years, the challenge to learn and adapt as a master student in a reputed university like TU Delft was not a cakewalk. However, a huge part of this process was accomplished with my constant support system back in India and in Netherlands. I would like to thank my parents – Amma and Acha who gave me the strength, guidance and confidence to uproot myself from India and start a new chapter in a completely different country. Thank you chechi, for being my sister and one of my closest friends from the time I can remember. My friends back home in Cochin – Arjun, Alka and Miriam who have been the closest to me from my high school days were always a call away for me to share and speak about any issues that I had.

Friendships that crossed boundaries are one my priced possessions in Netherlands. Becoming best friends with your assignment partner and sharing both formal and informal discussions with them is indeed rare. I would like to thank my best friend Sofia for her constant support, presence and patience throughout the span of the last two years. At the beginning of my life in Delft, I found a home away from home "33F, Aan Het Verlaat". My housemates here started out friends and then became my family in Europe. I would like to thank each of them for their constant support and presence. My classmates from T&P – Neeraj and Raunaq were also a part of my close friend's circle and I would like thank them for all the support they have extended to me over these two years.

I'm looking forward to start the next phase in my life in the field of transport planning and I hope my professional accomplishments contribute to researchers and professional worldwide.

Aswin Menon Nandakumar
Voorburg, December 2021

Executive Summary

Over the past few decades transport authorities globally have resorted to transport models for testing policy interventions and simulating the results as part of ex-ante analysis. To cater to different application requirements a large number of transport models are developed/under-development around the globe. Departing from the traditional four step process in traffic modeling, conventional assignment of traffic occurs under two main classifications – **Static** and **Dynamic**. While static assignment occurs with aggregated time-invariant interactions of traffic demand and supply, Dynamic Traffic Assignment (DTA) models seek to provide a detailed method to mimic the interaction between travel choices, traffic flows, and travel time measures in a temporally coherent manner (Chiu, et al., 2011). The dynamic representation of traffic has proved to be more accurate when compared to their static counterparts (Peeta & Ziliaskopoulos, 2001). However, due to limitation of hardware and software capabilities, a feasible simulation run using a DTA model is still under development.

Within the domain of DTA models, **Macroscopic representation of traffic – Macroscopic DTA's** takes place at an aggregate level, departing from the classical traffic flow theories. They simulate traffic analogous to the flow of fluids or gases. Due to the aggregation, these models run way faster than their microscopic counterpart, which gives them a clear advantage from the perspective of network scalability (Ferrara, Sacone, & Siri, 2018).

However, there exists another challenge with any model user to use the right model for the right application. In this regard there is an overarching emphasis on the expertise of the model user to select the correct model. Literature on evaluating Macroscopic DTA's were found to be scarce. Those articles which exist describe mainly their classification schemes - (Peeta & Ziliaskopoulos, 2001), (Chiu, et al., 2011) etc.; or are based on specific applications - (Flügel, Flötteröd, Kwong, & Steinsland, 2014), (Salgado, Jolovic, Martin, & Aldrete, 2016), (US Department of Transportation, 2004) etc. However, based on interviews with experts in traffic modeling, it is understood that evaluation of traffic models is subjective with multiple perspectives for sensitivity.

The research aims to strike this research gap, by developing a framework for evaluating Macroscopic DTA models. The research project is performed as part of Master Thesis of the author in collaboration with experts in traffic modeling at Dat. Mobility, Deventer and Delft University of Technology, Delft. Departing from this research context, the primary research question is formulated as described below:

How to compare Macroscopic Dynamic Traffic Assignment Models based on their performance under various evaluation themes?

Steered by this objective, the research involved the design, development and validation of a framework for evaluation called **EMMa** – **E**valuation **M**odel for **Ma**croscopic DTA's. The framework design is as shown in Figure A. EMMa is governed by four dimensions

1.) **The Measures of Performance (MoPs) and its type**: 7 sub-categories - *Conceptual Validation, Model robustness, Applicability, Tractability, Integration of Network Hierarchies - Urban and Motorway roads, Computational efficiency and Usability*

2.) **The Model User Type**: *Policy Maker, Mobility Consultant, Scientific Researcher, Model Developer*

3.) **The Application Planning Horizon**: *Strategic, Tactical and Operational Planning*

4.) **The DTA models** in itself applied for evaluation.

The objective evaluation of the DTA's is performed through MoPs. The scores of the MoPs is obtained by conducting a series of tests on theoretical and real-world large-scale traffic networks for the DTA models. On the basis of the measurement type, the MoP scores are quantitative, qualitative or binary.

Motivated from the structure of a Multi-Criteria Decision Analysis, the subjective side of EMMa showcases the differences in importance associated with model features which varies from model users to application domains. The weights for these subjective dimensions are obtained by conducting surveys and interviews with traffic experts across the four model user types.

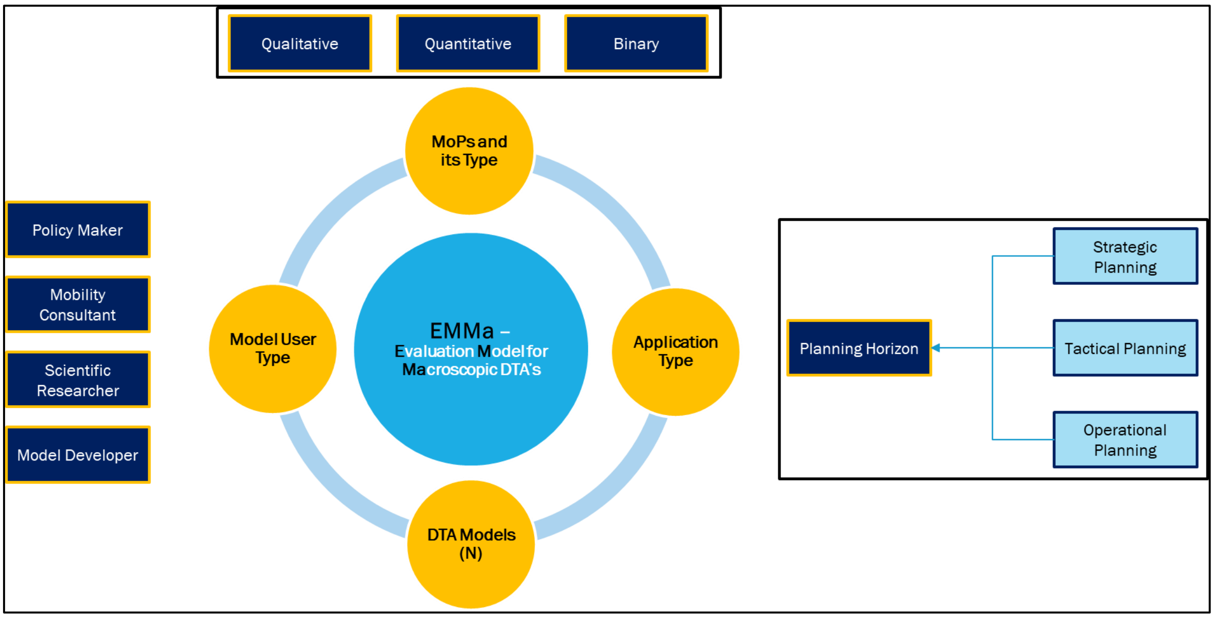


Figure A: Dimensions of EMMa

The macroscopic DTA models used for comparison and application of EMMa are: the MARPLE (Model for Assignment and Regional Policy Evaluation), StreamLine: MaDAM (Macroscopic Dynamic Assignment Model) and StreamLine: eGLTM (event based Generalized Link Transmission Model). The models are selected on the basis of availability, access to software and variability observed in the modeling choices such as link propagation, junction modeling, route choice models used for achieving user equilibria etc.

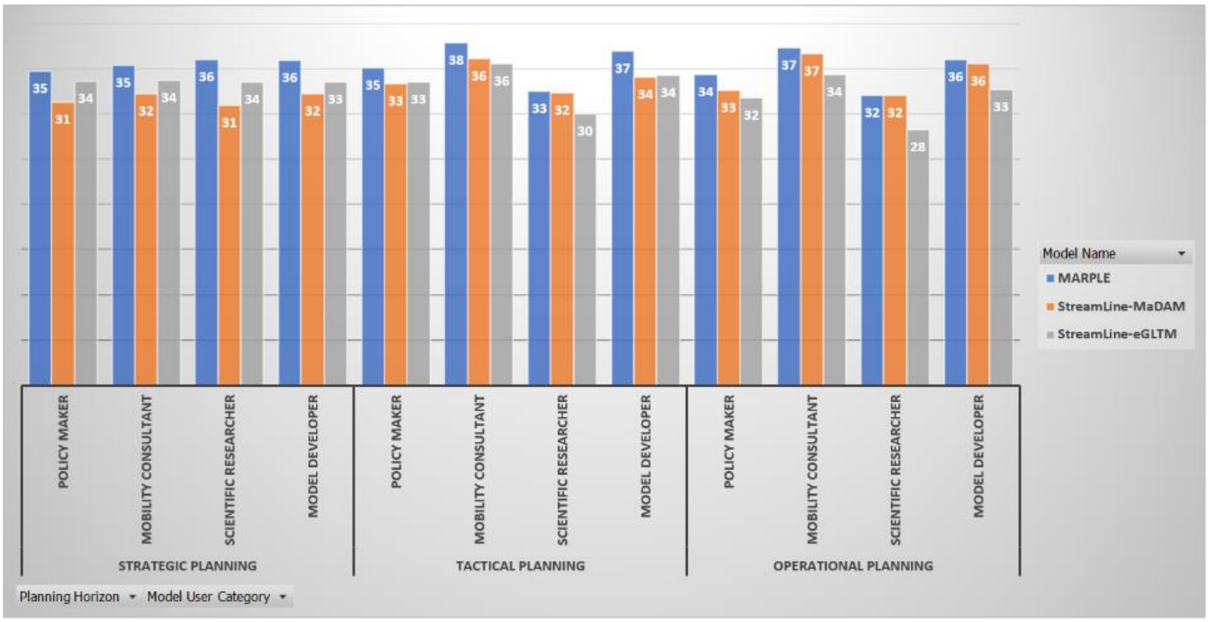


Figure B: Final Results of Evaluation

The final results of evaluation are summarized in Figure B. The variation in results with respect to Model User perspective and Application planning perspective can be observed here. For Strategic Planning, both MARPLE and StreamLine: eGLTM were the clear achievers, with former performing slightly better. The strength of both these models was in the MoP category of Model Robustness, which re-validates the importance of a stable state of equilibrium for large-scale strategic planning application. Furthermore, both the models showcased fast and efficient computation capabilities. As the time horizons of application became smaller as is the case with Tactical and Operational planning, the final score for StreamLine: MaDAM improved substantially across all model users. The strength of StreamLine: MaDAM were mainly in its accuracy involved in link-level propagation and queuing. The second-order propagation model in MaDAM further boosted its score in modeling propagation in urban and non-urban links adequately. However, the computational efficiency of the network loading algorithm in MaDAM was poor. This hindered its achievement as the best model especially in Operational Planning applications, where the need for high-speed computation was of utmost importance across most model users.

The ability of a DTA model or any transport model for that matter, is to simulate the behavior of a transport system adequately within a virtual environment, which acts as a safe haven for trails and experiments. It becomes clear to any model user or a developer that an ideal model does not exist but rather serves as a tool for decision-making for the problem at hand on the basis of some theoretical assumptions. Thus, the choice of the model is a key criterion in finding solutions to the problem. The framework EMMa thus serves as a model for macroscopic DTA models to help the modeler to choose the correct model. From the application of EMMa to the three models selected for this research, the fundamental trade-off between model complexity and computational speed was clearly visible from the results. MARPLE owing to high-speed computation capabilities and faster achievement of a stable equilibrium state proved to achieve Rank No.1 across most model user categories and application horizons. This can be interpreted that, model users in general valued these model characteristics over complexity of results (through various complex features of the model as is the case with MaDAM). However, we observe variations across model users, which validates our original hypothesis that the right choice of a model primary depends on the person using it and the application it is deployed for.

Table of Contents

List of Figures

List of Tables

1. Introduction and Background

Transport planning involves the composition of transportation systems which includes infrastructure for the different vehicle types, such that the travel demand can be accommodated safely and efficiently. For over several decades now, the role of transport planning has been paramount for proactively identifying, tackling, and preventing the problems of transport systems such as congestion, difficulty in pedestrian mobility, environmental impacts, etc. Before the implementation of solutions to tackle these issues, they are simulated and tested in a risk-free environment through transport models. The results from these models support decision-making. Evolution of IT (Information Technology) and associated hardware helps in the realization of new infrastructure concepts (eg. Intelligent Transport System- ITS), mobility systems (eg. autonomous vehicles, Demand Response Transit – DRT, etc), electronic payment systems (smart cards, app-based tickets), etc., within a transport model.

Experts argue that the main limitation involved in decision-making is the technical proficiency of transport professionals and the knowledge of theoretically sound modeling techniques with their feasible software implementations (Ortúzar & Willumsen, 2011). A model is a simplified representation of a part of the real-world system of interest, which focuses on certain elements considered important from a particular point of view (Ortúzar & Willumsen, 2011). From this definition, it can be understood that the success of a model depends on the adequacy of its application domain and the problem it addresses. The feature offered by such models to mimic and experiment with policy-based scenarios is where its key strength lies. Transport models serve as a tool to forecast the outcome of their decisions, thereby serving as an aid for decision making, proposing new legislations, or approving new infrastructure projects. This is a fairly recent trend as the widespread adoption of transport models only started in the middle of the twentieth century (Raadsen M. , 2018). Along with the advanced computing power of the digital hardware, the time required to run the model to simulate results also became an important factor for the adoption of a particular model.

1.1. Traffic Assignment Models

Traffic assignment models in particular deal with the interaction between traffic demand and supply. They are employed to simulate traffic flows on a network. The traffic assignment model predicts the network flows that are associated with future planning scenarios and generates estimates of the link travel times and related attributes that constitute the basis for benefits estimation and other assessment criteria.

The demand model component is responsible for estimating the traffic demand based on traveler preferences, socio-economic data, etc. These models use an input matrix of vehicular flows that represents the volume of traffic between origin and destination (O-D) pairs. The supply model describes the physical traffic network, which consists of the network topology, link characteristics, link performance functions, etc. The flows for each O-D pair are loaded onto the network based on the travel time or impedance of the alternative paths that could be chosen. The interaction between demand and supply, results in the demand being distributed across the network paths. This process is termed the traffic assignment. With the addition of time-varying demand-supply interactions to traffic models in a behaviorally sound approach, Dynamic Traffic Assignment (DTA) Models are created.

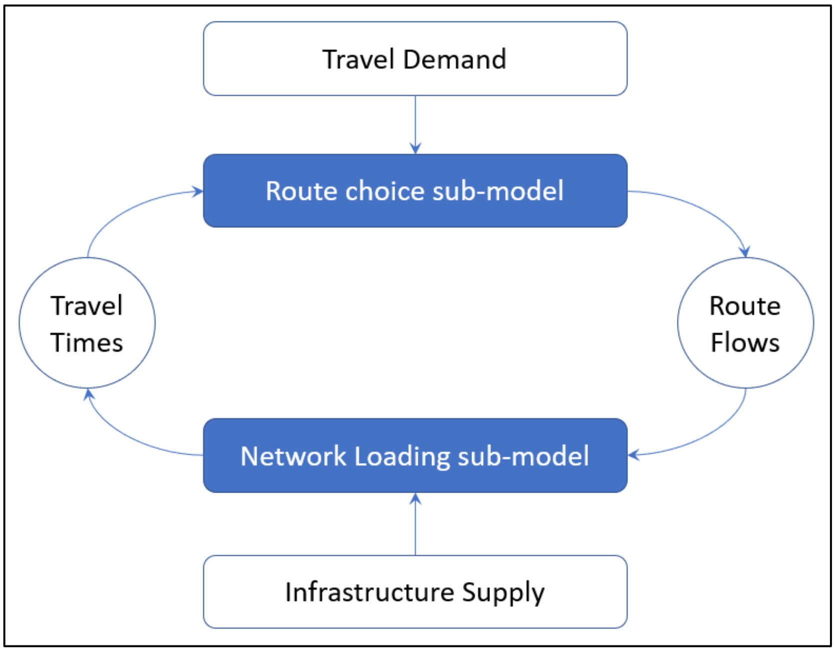


Figure 1: Interaction between travel demand and infrastructure supply adapted from (Bliemer M. C., et al., 2017).

A theoretical interaction diagram is shown in Figure 1. For every iteration, the travel demand acts as an input to the route choice sub-model. The route flows are derived from the choice models, route proportions, and travel demand. In some DTA models, the route choice proportions are updated during the simulation run. In most models, the route fraction calculation occurs only at the start of the time period, for each iteration of the simulation. The route flows are assigned by the network loading model yielding link flows, densities and speed. The link travel times are derived from these link speeds. The link travel times lead to route travel times which are sent as feedback to the route choice module. The equilibrium conditions are re-checked by evaluating the duality gap value and comparing it to a given threshold, which acts as the primary stop criterion. In the case of non-convergence, the secondary stop criterion (a maximum number of iterations) is employed. The route choice proportions are re-calculated based on the updated route travel times and the interaction loop is continued.

In the context of DTAs, based on granularity, three main variants exist: Microscopic, Macroscopic and Mesoscopic DTA models. **Microscopic models** capture the dynamics of all vehicles and their interactions are represented at the finest level of detail. Due to an individual perspective, in most cases, these models are computationally expensive, especially for a large-scale network. **Macroscopic models** on the other hand represent traffic at an aggregate level. They simulate traffic analogous to the flow of fluids or gases. The dynamics of traffic in such a case are described using aggregate variables such as density, mean speed, and flow. Due to the aggregation, these models run way faster than their microscopic counterparts, which gives them a clear advantage from the perspective of network scalability (Ferrara, Sacone, & Siri, 2018). Between aggregate and microscopic representation of traffic lies **Mesoscopic models**. They do not distinguish individual vehicles but mimic and model the heterogeneity of the driver's choices in probabilistic terms (Ferrara, Sacone, & Siri, 2018).

1.2. Thesis Contribution and Research Context

The usefulness of a transport model will strongly depend on the application, the experience of the modeler, and the ability of the model to represent the problem at hand, feasibly. With the variations involved in the modeling mechanisms and the growing number of simulation software's, the model user is faced with the challenge to use the right model for the right application. This involves evaluating the available models at hand under various modeling properties i.e., a multi-dimensional framework to compare and rank the models. The framework should be multi-disciplinary to incorporate the various measures of model performance along with the intended application and the type of model user.

For the scope of this research, such an evaluation is targeted on Macroscopic DTA models. Thus, the research will involve the creation of a framework for evaluating and comparing Macroscopic DTA models. The scope of this research is limited to such models due to constraints of time and access to the models. The macroscopic DTA models used for comparison and application of the evaluation framework for this thesis project are: the MARPLE (Model for Assignment and Regional Policy Evaluation), StreamLine: MaDAM (Macroscopic Dynamic Assignment Model), and StreamLine: eGLTM (event-based Generalized Link Transmission Model). The selection of the models is motivated in Section 2.3. Essentially the framework will aid the model user in choosing a Macroscopic DTA model. As part of the research, three key deliverables are provided: 1) The thesis report, 2) A working model of evaluation framework 3) A set of theoretical test networks (developed partly by the researcher and partly from existing sources) which can isolate and test an individual sub-module of the DTA.

1.2.1. Research Objectives

The main research objectives set for this thesis is as described below

- Identification of key modeling properties of Macroscopic DTA models from literature as Measures of Performance (MoPs) to be used for model evaluation.

- The creation of a framework and in effect a multi-dimensional tool for evaluating macroscopic DTA models based on relevant MoPs, application domains, and model user types.

- A comprehensive study on DTA models – MARPLE, MaDAM, and eGLTM. A literature review on earlier work along with a focus on the theoretical background on the models. The focus will be on the Dynamic Network Loading (DNL) sub-component.

- Application of the framework across the three models on theoretical test networks and a large-scale real-world traffic network for quantitative and qualitative evaluation.

1.2.2. Research Question

Based on the above objectives the following research question and sub-questions are framed as shown below:

Primary Research Question: How to compare Macroscopic Dynamic Traffic Assignment Models based on their performance under various evaluation themes?

Sub Question-1: How representative are the DTA models chosen for conducting the current research? what are their strengths and weaknesses? (Qualitative classification); - Section 2.3.3

Sub Question-2: What are the measures of performances that will be used to evaluate under each application scenario – strategic/tactical/operational, and with different model user perspectives? Section 3.2

Sub Question-3: How do the models score and rank the models based on the evaluation criterion? Section 3.5

As seen in the above main research question and the sub-questions, the research aligns with the formation of an evaluation framework that will be used to compare the various DTA models, with a focus on its variability with applications and model user type. The sub-questions direct some sub-tasks which are required to formulate the evaluation framework and apply the same to rank the models.

1.2.3. Research Methodology

The research methodology is formulated based on the research objectives and the direction is aligned so that the output of one step becomes an input to the subsequent step. The approach in general is linear with additional inputs which are specific to certain parts of the research. The detailed research methodology is shown in Figure 2. The text provided in the blue box represents the tasks that are executed during the research and the text in red are the output of the tasks.

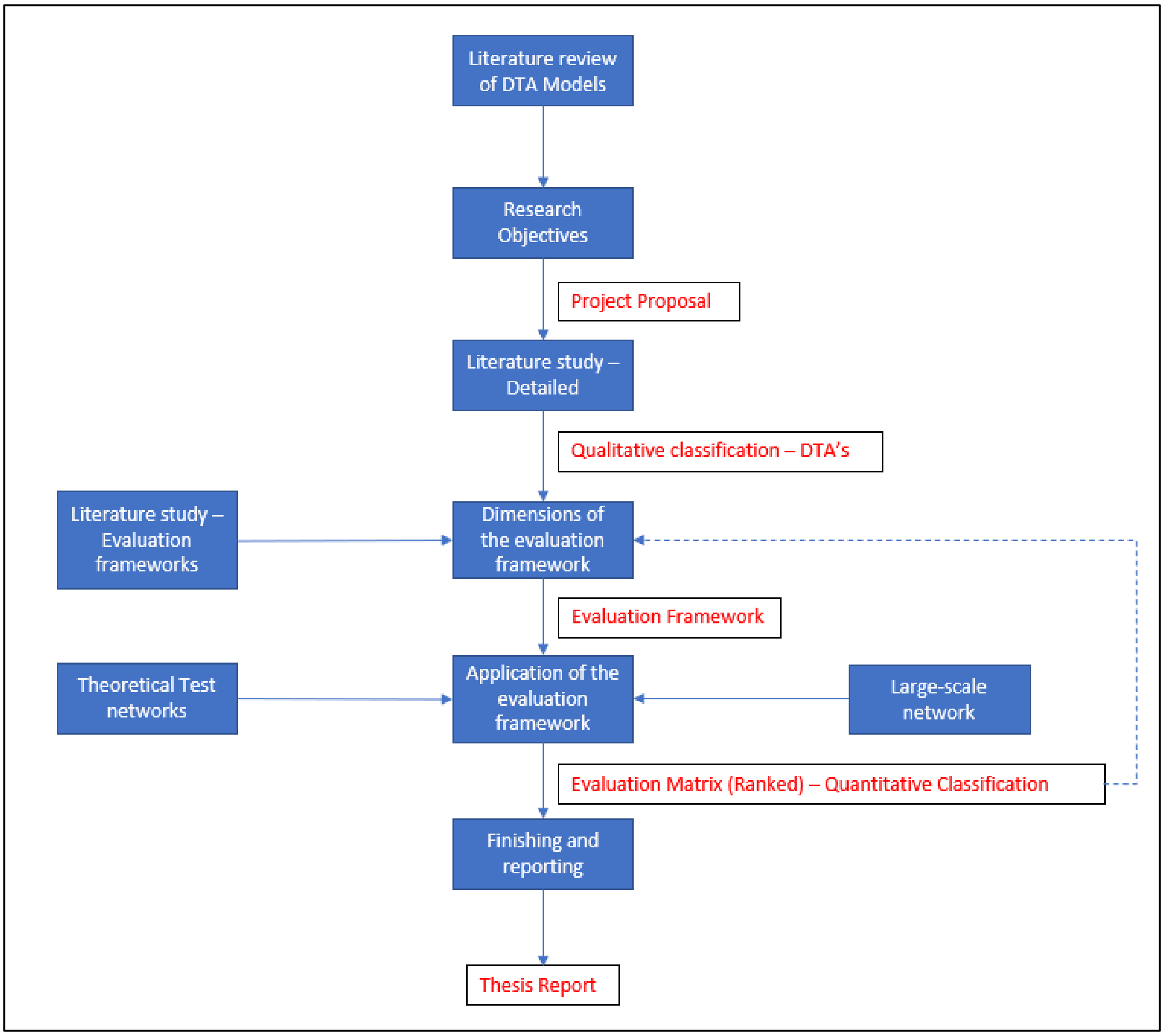


Figure 2: Research Methodology

The first part of the research starts with the **literature study** and the theoretical background of the DTA models in general (See Section 2.1). Based on the initial research, scope definition in the form of **research objectives**, research questions, thesis proposal, and the workplan is aligned.

A **detailed literature study** on the three models is conducted as the next step which is used to formulate a classification table with the key modeling properties (See Section 2.2). Literature study also motivates the selection of these DTA models used in this research, as summarized in the qualitative comparison

(Section 2.3.3). Aspects from this classification table are further used in the evaluation framework during scoring. The next step involved the formation of the framework for performing the evaluation. The start point of the framework is the classification table. However, excerpts and ideas from existing literature are used as motivation to design the framework and the consequent evaluation tool (Section 2.4). This step paves the way to the **formation of framework dimensions** and subsequently the framework itself (See Section 3.1 and Section 3.5).

The framework is then applied for evaluating the DTA models selected for this research. Based on the MoPs, each model is scored and an evaluation matrix is created (Section 4.1). The matrix provides suggestions on the application of the traffic models and discusses the strengths and weaknesses of each of the DTA's, tailored to fit a specific model user perspective. The thesis covers extensively the theoretical validation of the MoP through the results of the simulation models. Validation of the models through empirical data is excluded from the scope, due to constraints of time and data availability. The framework dimensions are also reassessed based on the quantitative evaluation. This is shown through the feedback loop in the methodology diagram in Figure 2. The final step in the research process is the finishing and reporting of the results along with the submission of the thesis report.

1.3. Organization of the Report

The Thesis report is organized as follows, it contains Five Chapters, where Chapters 1 and 5 serve as Introduction and conclusion, and the three parts within form the core. The structure of the report is linear to the extent of information flow. **Chapter 1: Introduction and Background**, introduces the reader to the premise of this research project with a briefing on Traffic Models and the need for evaluating them. As a consequence of this need, the research gap and relevance are described with details in the research context and methodology.

Chapter 2: Literature Overview is divided into 3 parts, with a focus on Macroscopic DTA algorithms, the models used in the current project, and the theoretical backing for forming the evaluation framework developed in this thesis. **Chapter 3: Methodology**, provides the reader with a detailed explanation of the various attributes and features of the evaluation framework- EMMa (Evaluation Model for Macroscopic DTAs). The methodology also touches upon some literature references used for developing EMMa. A briefing on the theoretical test networks used for the project is provided in this section. The scoring, ranking, and weighing of the MoPs is another aspect described in this chapter.

The Methodology concludes with the final form of EMMa and paves the way into the application of the framework for the three traffic models identified for the current research, which is described in **Chapter 4: Results & Discussion**. This chapter showcases the results for all three models for the same set of theoretical test networks. Furthermore, some qualitative MoPs are scored for the three models, which can directly be identified by understanding the theory behind the traffic models. The scores for the MoPs are motivated either through literature backup or through a series of simulation tests as documented comprehensively in this section.

The thesis report completes with **Chapter 5: Conclusions and Future Recommendations**. The chapter discusses in detail the final remarks from the results section. The design of the framework was subject to a series of shaping and fine-tuning, which has opened up numerous research topics which can be an extension of this thesis. This is the final focus of the thesis report under the section – Future for EMMa. The section-wise brief overview is as provided in Figure 3.

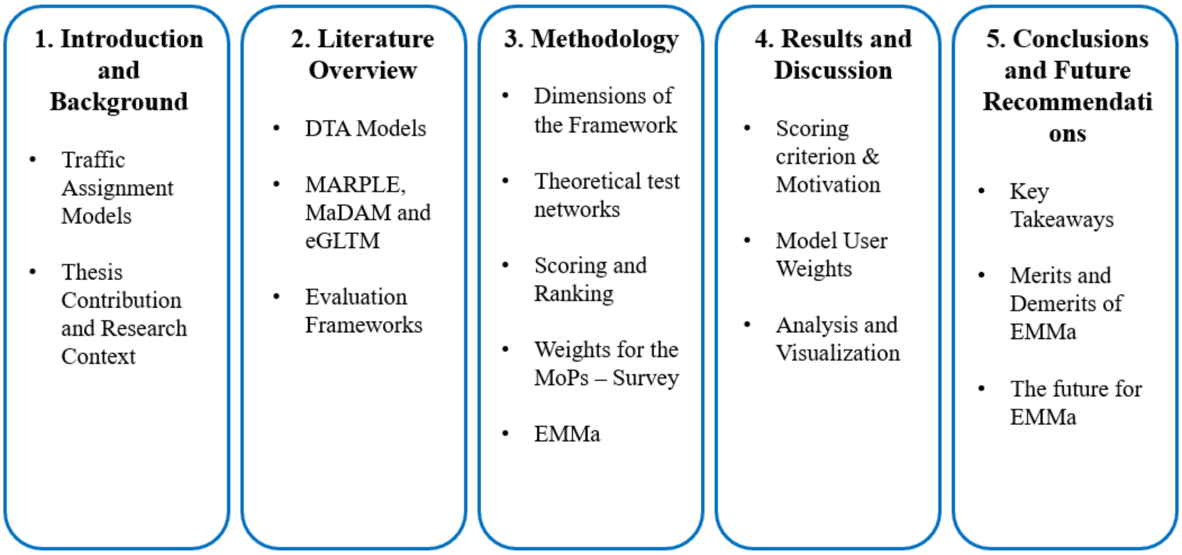


Figure 3:Thesis Outline

2. Literature Overview

The current research involves evaluating various simulation-based Traffic Assignment models with a special focus on Macroscopic DTA models. When compared to their microscopic counterparts, macroscopic DTA's can prevent statistical noise during simulation runs (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019). To explain this further, in microscopic simulation, vehicles are selected at random, which departs from their origins at specific time instances. This causes statistical noise, which causes issues when comparing outcomes of two different simulation runs, of the same scenario. Furthermore, macroscopic models can provide deterministic outcomes (results are floating-point numbers instead of integers), in comparison to microscopic models. This is expected to improve the convergence since it allows for solution search between routes on an OD pair using any number and not just integers. The above-mentioned advantages when coupled with a lower computation time give macroscopic DTA models a clear advantage for a wide range of applications.

The literature study performed as part of this research will focus on mainly three aspects.

- A brief overview of DTA modeling, the modeling components/theories, and classifications present in existing literature, with a focus on macroscopic DTAs, will be the first part (see Section 2.1). It may be noted that the review provided below is not exhaustive and highlights the most common models within each category.

- The literature review about simulation-based Macroscopic DTA models (see Section 2.2) and its motivation for the choice of the three models analyzed in this thesis: MARPLE, MaDAM, and eGLTM (Section 2.3)

- In addition, the current evaluation framework and its dimensions are greatly inspired by other evaluation studies for transport models and projects. This forms the third part of the literature review (Section 2.4).

2.1. Dynamic Traffic Assignment Models

The static representation of traffic has notable advantages in its theoretical tractability of mathematical properties (e.g., existence and uniqueness of equilibrium), which can be obtained relatively easily. Furthermore, the ease of finding an approximate solution for the problem, in a static setting comes at feasible computational complexity. At the peak of its development, the main application of static models was for strategic planning applications such as large-capacity expansion projects (Chiu, et al., 2011). However, a pure static model, by definition, is unable to mimic the evolution of traffic flows over time and consequently the dynamic variation in traffic behavior. Thus, pure static assignment becomes badly suited to analyze either traffic congestion effects at a fine-grained temporal level or many of the solutions that can be taken to address congestion.

In this regard, DTA models have witnessed a huge research focus since the seminal work of Merchant and Nemhauser (Nemhauser1 & Merchant, 1978), (Nemhauser2 & Merchant, 1978). (Peeta & Ziliaskopoulos, 2001) provides an extensive overview of DTA models. There is a heightened interest in DTA, especially for the development of methods that can be used for large-scale real-time and planning applications. One common feature of these models is that they depart from the standard static assignment assumption of stationary demand during a single time period. However, the advantage of DTA comes from its ability to use the merits of a static assignment for a time-varying demand and network loading.

Modeling of queues is done more realistically in dynamic models due to strict adherence to link capacity and storage constraints. As a consequence, the route choices behave differently when conducting equilibrium assignments. Recent developments show a trend towards a mixture of static macroscopic models with properties typically only used in DTA models, such as the incorporation of strict capacity or even strict capacity and storage constraints. An example of such a model is STAQ (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019).

Difference from Static Assignment Techniques

Static assignment models are used extensively due to their ability to provide approximate solutions through a User Equilibrium (UE) approach. To retain this advantage in the case of a dynamic model, the assumptions need to be extended in two ways (Chiu, et al., 2011). In a static model, route choice is assumed to be pre-trip, which may be deterministic (shortest travel cost) or stochastic (perceived shortest travel cost). As a primary extension, in the case of a DTA, this is done by recognizing time-dependent travel times/costs in links. Travelers are assumed to know or anticipate future travel conditions along the journey (through learning from the past trials) and, in choosing an O-D route, they are assumed to minimize the O-D travel time that they will experience. This is depended on when they arrive at the various links along a route and on the travel times/costs that are present on the links at those specific future times.

Building on the lines of variation in generalized route cost over the different time periods for a particular OD pair, DTA recognizes the need to ascertain user equilibrium conditions for all travelers at a particular departure time. To explain further, in a dynamic approach, the user equilibrium condition of equal generalized travel costs on used routes applies only to travelers who are assumed to depart at the same time between a particular O-D pair. In a DTA model, this can be further extended by incorporating a choice of departure time for the travelers simultaneously. Such a feature can be used to analyze phenomenon such as peak spreading in response to temporal variation in congestion, time-varying tolls, etc. However, the framework and models proposed/used in this research will not be considering this sub-component.

2.2. Towards Simulation-Based Macroscopic DTA

Fundamentally, all DTA models aim to adequately represent traffic realism and human behavior. This objective is further constrained by the time dependency and the randomness in the system inputs. The ability of the model to effectively capture realistic traffic dynamics comes at the price of reduced theoretical tractability. This inherent trade-off in DTA has created a distinction of approaches– analytical and simulation-based. The analytical approach is further classified into three different categories: mathematical programming, optimal control, and variational inequalities approach. The reader is directed to (Peeta & Ziliaskopoulos, 2001) for an extensive read about the various approaches and the literature present in this field.

One of the primary constraints related to the practical application of the analytical DTA model is scalability. Therefore, for the scope of this research, simulation-based DTA models will be considered and further studied, as their analytical counterparts cannot be applied even to the smallest real-world traffic networks. In simulation-based models, a traffic simulator is used to replicate the complex traffic flow dynamics. In an analytical formulation, critical constraints that describe the traffic flow propagation and Spatio-temporal interactions, such as the link-path incidence relationships, flow conservation, and vehicular movements, are addressed through simulation instead of analytical evaluation while solving the problem. Hence, a simulation-based primarily refers to the solution methodology rather than the problem formulation (Peeta & Ziliaskopoulos, 2001).

One of the main problems associated with simulation-based models in comparison to analytical is the absence of a strong theoretical guarantee of properties such as existence, uniqueness, and stability of solutions which are usually tenable only through compromises in mimicking the traffic phenomenon and restrictive assumptions on driver behavior (Peeta & Ziliaskopoulos, 2001). Simulation-based models strive to achieve adequacy (including real-world utility) over accuracy. From a practical and application perspective, the requirement of a strong mathematical background (uniqueness and/or global optimality of solutions) may not be a key requirement as long as the conditions on the road are satisfactory. Furthermore, the flexibility offered through simulation models to obtain realistic traffic flows has enhanced the acceptability of such models in the context of real-world deployment (Mahmassani, Chiu, Chang, Peeta, & Ziliaskopoulos, 1998) (Ben-Akiva, M., Koutsopoulos, & Mishalani, 1998). Due to the substantial computational burden associated with the use of a simulator, the choice of granularity (microscopic, macroscopic, or mesoscopic) has significant implications for the tractability of such models. In other words, such a classification is based on the level of detail.

The classification schemes described in this research is adopted from (Hoogendoorn & Bovy, 2001) and (Kessels, Lint, Vuik, & Hoogendoorn, 2014). An aggregate representation of traffic that evolves over space and time is the objective for any macroscopic DTA. Therefore, the primary distinction occurs in the type of representation over space and time: continuous and discrete traffic models. **Continuous** models involve the consideration of space and time as continuous variables and, consequently, the dynamics of the system are represented with differential equations. **Discrete models** involve the discretization of space and time. Specifically, a road traffic system is divided into a set of road portions with finite length, and the time horizon is subdivided into a given number of time intervals.

Macroscopic models are further distinguished based on several state variables: **first order**, **second order,** and **higher order**. For the scope of this literature review and due to the limited availability of existing research in the field, we will be restricted to first order and second order models. The most basic of the traffic modeling is first order modeling. In this case, a fundamental diagram is assumed, and hence there is a relationship. In second order models, additional terms are introduced in the relationship which incorporates speed dynamics such as acceleration, deceleration, inertia effects, etc. The following sections will cover this in detail.

2.2.1. Continuous Case: First Order Models

For any generic location x and time t on a road, the variables considered in continuous macroscopic traffic models are:

- $\rho(x,t)$ – traffic density (veh/km)
- $v(x,t)$ – average speed (km/h)
- $q(x,t)$ – traffic flow (veh/h)

The basis of every macroscopic model are two important relations, the first one being the hydro-dynamic equation as stated below:

$$q(x,t) = \rho(x,t).v(x,t) \qquad (1)$$

The second relation is the continuity equation or conservation equation derived directly from the conservation law of vehicle flows as stated below:

$$\frac{\partial\rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial t} = 0 \qquad (2)$$

All the continuous macroscopic traffic models are based on (1) and (2) and differ for the other equations which relate the variables $\rho(x,t)$, $v(x,t)$ and $q(x,t)$. The theoretical relation between density and flow in steady-state conditions is the Fundamental diagram. This is a relation $Q(\rho(x,t))$, which has to satisfy the following criterion:

$$Q(0) = 0, Q(\rho^{max}) = 0, \frac{dQ(\rho)}{d\rho}\bigg|_{\rho\, =\, \rho^{cr}} = 0 \tag{3}$$

where, ρ^{cr} is the critical density (veh/km), ρ^{max} is the jam density (veh/km) and q^{max} is the capacity in (veh/hr). There exists a wide range of fundamental diagrams of different shapes to represent steady-state relationships.

The first macroscopic model was developed by (Lighthill & Whitham, 1955) and (Richards P., 1956), famously known as the Lighthill-Whitham-Richards (LWR) model. The model captures the dynamics of a single variable – Traffic Density. For the variations in the LWR model, associated literature and the mathematical formulations, interested readers are referred to (Ferrara, Sacone, & Siri, 2018).

2.2.2. Discrete Case: First Order Models

Discretization of space and time into finite units is the fundamental concept of discrete first order models. In most cases, the continuous LWR models is discretized based on a various numerical method. According to these methods, the road space is divided into portions of finite length, time is split into intervals of equal duration, and the partial differential equation of the LWR model is transformed into a finite-difference equation.

The most famous LWR based model, which is discretized is the Cell-Transmission Model (CTM) given by Daganzo in (Daganzo C. , 1993) and (Daganzo C. , 1994). The initial model was for a one-way road without any intermediate entrances and exits. This is was further extended to a road network case with a three-legged junction in (Daganzo C. , 1995). Traffic interaction features such as on-ramps, off-ramps freeway junction features, etc. were incorporated in this model. According to CTM, the space is discretized into units called cells. The boundary between two cells is governed by two quantities: *Sending function*, which is dependent on the density before the junction, and the *Receiving function*, depending on the density downstream of the junction. A detailed overview of the various extensions of CTM is provided in (Ferrara, Sacone, & Siri, 2018).

Of the reviewed articles for the discretization of first-order models it is interesting to note the detailed overview given by Lebaque in (Lebacque, 1996). In this paper, it is shown that the CTM corresponds to the application of the Godunov scheme to the LWR model. The Gudonov scheme is a numerical method introduced in (Godunov, 1959). By applying the Gudonov scheme, a condition for the space discretization L and time discretization T is also derived which can be expressed as

$$T \max_{\rho\epsilon[0,\rho^{max}]} \left|\frac{dQ(\rho)}{d\rho}\right| \leq L \tag{4}$$

The expression states that for the product of the discrete time step (T) and the max speed propagated in the model for all the densities in the fundamental diagram, the distance covered should be either less than or equal to the discrete cell length (L), input in the model. This ensures that at any given time step, no vehicle travels in two or more different cells. Furthermore, Lebaque introduces the terminology demand and supply for the sending and receiving functions in the CTM.

Among the various discretization's available for the LWR model, an important variant is the Link Transmission Model (LTM) proposed for the first time by (Yperman, Logghe, & Immers, 2005). In the LTM, traffic propagation is represented as the cumulative number of vehicles that pass through the beginning and end of a link per time step. In other words, this can be translated as numerical calculation occurring only at the link boundaries as opposed to each cell boundary in a CTM. This forms a clear computational advantage for the LTM over the CTM. Moreover, errors related to averaging of discrete space can be avoided through LTM. The reader is referred to (Papageorgiou M. , 1998) and more recently (Bliemer M. C., et al., 2017) for further discussions on first-order macroscopic models.

2.2.3. Continuous Case: Second Order Models

First order models such as LWR and its extension in the form of first order CTM and LTM models, present some limitations in terms of dynamics of traffic representation. For instance, it does not contain any inertial effects, as it assumes vehicles adjust their speed instantaneously. This can lead to unrealistically high accelerations or decelerations of vehicles. Another phenomenon not considered by such models is capacity drop, observed in real-world traffic networks. The model systematically predicts that the output flow from a congested area is equal to the capacity flow if the road downstream is not congested which is not the case in real-world.

To tackle the limitations of first-order model, second order models came into development. Besides considering the dynamics of the traffic density, these models explicitly introduce a dynamic equation for the mean speed. The variation in the speed equation incorporates hysteresis observed in real-world traffic movement. The first continuous second-order traffic flow model was proposed by (Payne, 1971) and (Whitham, 1974) and is generally known as the Payne-Whitham (PW) Model. For comparison, this literature review will briefly summarize the mathematics behind the PW model. Interested readers are referred to books dedicated to continuous traffic models such as (Garavello & Piccoli, 2016) and (Garavello, Han, & Piccoli, 2006).

Similar to first order models, the PW model is also governed by the hydrodynamic equation and the continuity equation given in (1) and (2). However, the model is also coupled with a partial differential equation describing the dynamics of mean speed. This expression is as stated below:

$$\frac{\partial v(x,t)}{\partial t} + \underbrace{v(x,t)\frac{\partial v(x,t)}{\partial x}}_{Convection\ term} = \underbrace{\frac{1}{\tau}\left[V\big(\rho(x,t)\big) - v(x,t)\right]}_{Relaxation\ term} + \underbrace{\frac{1}{2\tau\rho(x,t)}\frac{dV(\rho)}{d\rho}\frac{\partial\rho(x,t)}{\partial x}}_{Anticipation\ term} \qquad (5)$$

The <u>convection term</u> describes the influence of the upstream speed on the speed downstream. This term ensures that the vehicles travelling along the freeway do not adjust their speed instantaneously. To provide an example let us consider the case in which vehicles are travelling very fast and need to decrease their speed to adapt to the lower downstream speed. The vehicles gradually reduce their speed which implies that a higher upstream speed tends to increase the vehicle speeds downstream and the opposite in case of a lower upstream speed.

The <u>relaxation term</u> models the fact that all vehicles adjust to the speed in steady state represented by $V\big(\rho(x,t)\big)$. The constant τ represents the speed relaxation time and is related to the reaction time of the drivers.

The <u>anticipation term</u> describes the capability of the drivers to look ahead and adjust their speed according to the speed of the vehicles downstream. This implies that this term models the speed adjustment of vehicles to a value compatible with the density downstream.

The thought process behind the creation of the PW model was the analogy of vehicles to that of fluids. However, it can be understood that the behavior of traffic is anisotropic, as we are dealing with human drivers with a personality, and inconsistencies in behaviors such as negative speeds cannot be tolerated while modeling a real-world traffic system. This led to the development of the ARZ model given in

(Aw & Rascle, 2000) (Zhang, 2002). Interested readers are referred to (Ferrara, Sacone, & Siri, 2018) for a detailed explanation of the model along with its extension to road traffic networks.

2.2.4. Discrete Case: Second Order Models

The first discretized version of the PW model came in the form of an application in Paris, through the METANET simulation-based model (Papageorgiou, Blosseville, & Hadj-Salem, 1989) (Papageorgiou M. , 1990) (Messmer & Papageorgiou, 1990). The METANET acronym stands for '*Modèled' Écoulement de Trafic sur Autoroute NETworks*'. The model is used as an extension to the PW model for the application on a freeway network considering the discrete case of the space and time variables, along with some new terms to model the influence of on-ramp and off-ramp traffic flows.

Other than the convection, relaxation and anticipation a fourth term was introduced in (Papageorgiou, Blosseville, & Hadj-Salem, 1989). The deceleration caused in the mainstream due to lower speed traffic entering from the on-ramp is being modelled through the fourth term. This is especially relevant when the entering flows are high. Readers interested in other relevant METANET variants are referred to (Ferrara, Sacone, & Siri, 2018) for a detailed overview.

2.2.5. Multi Class Models: First Order and Second Order

Multi-class models have been developed to differentiate between various classes of vehicles travelling in the same road system. A multi-class macroscopic model, assumes that traffic behavior is represented by different traffic flows corresponding to different vehicle categories, whereas a single-class model assumes the whole traffic as a single homogenous fluid. The terminology of multi-class may refer to different mode types such as cars, trucks, public transport etc. or may refer to traveler types such as driving behavior, trip purpose etc. This is especially relevant while modeling ITS (Intelligent Transport System), where level of information for the driver plays a key role in traffic dynamics such as route choice. Intelligent vehicles can be characterized by those vehicles equipped with innovative technology enabling the exchange of data with other vehicles and the traffic infrastructure.

For first order models, the multi-class version can be categorized into those extensions of the LWR model and those for the CTM model, representing the continuous and the discrete cases. Similarly, in the case of second order models, there is extensive literature available for both the continuous case and the discrete case. However, the scope of this literature review is limited to single user class. Interested readers are again referred to (Ferrara, Sacone, & Siri, 2018) for an extensive overview on these models.

2.2.6. Classification based on Spatial Assumptions

Another classification scheme adopted from (Bliemer M. C., et al., 2017) based on assumptions in the spatial model is: capacity unrestrained, capacity restrained, capacity constrained, capacity and storage constrained. The authors provided this classification scheme, essentially for strategic transport application; however, we will explore the extension of this to tactical and operational applications in this research.

Capacity unrestrained models are those which allow traffic flow through their links, irrespective of the volume of the flow and the capacity of the link. In other words, they simulate free flow travel time in their links irrespective of the link volume. This approach is commonly used for an initial analysis which may involve finding potential paths, initializing more capable assignment types etc.

The second class of model is the capacity restrained type, which considers link capacities but to a certain extent. These models still allow the oversaturation of links with flows above the capacity. However, they enable these models to use link performance functions such as BPR functions (Bureau of Public Roads, 1964) or Akçelik functions (Akçelik R. , 1991) to divert the traffic from oversaturated links.

The link travel times are updated through these functions and as a consequence route travel time and route choices are altered per simulation. These models also guarantee the existence of unique optimal solutions for convergence thereby ensuring mathematical tractability. Due to this, there exists literature to find solution schemes for such models in strategic planning applications (Dial, 2006) (Bar-Gera, 2010), even though the results can be unrealistic mainly for congested conditions.

To overcome the infeasibility of capacity restrained model, a more capable alternative is the capacity constrained model. In such a model, the constraint imposed for restricting traffic flow in an oversaturated link is usually modelled in two ways:

- In earlier models, a side constraint was imposed which no longer allowed the flow to increase the capacity by employing a penalty function in the form of a Lagrange multiplier. These penalties induce a queuing delay as a result of congestion. Such a mathematical construct diverts the flows from congested to uncongested links. Examples of models which use such a multiplier are (Shahpar, Aashtiani, & Babazadeh, 2008) & (Larsson & Patriksson, 1995).

- Delay modeling based on actual demand on the links seemed more natural than a mathematical function. This need amongst others led to the development of models which incorporate vertical or horizontal queues. In a vertical (point) queue model, the queues are allowed to grow indefinitely before a bottleneck link, which is not the case in case of a horizontal queue model, which simulates the growth of the queue to adjacent links and showcase a spillback effect. The former model is less realistic, but if the queue does not grow beyond one link length, both the models give exactly the same result (Raadsen M. , 2018). Examples of dynamic (vertical) queue models include (Pang, Han, Ramadurai, & Ukkusuri, 2012) and (Smith, 1993).

Nodes in a transport network represent conflict points between two links, where vehicles interact which could also be a junction. In capacity constrained models, the available capacity at the exit links from a node need to be distributed among the incoming links. This allocation is the responsibility of the node model. The first macroscopic models only considered interactions at freeways such as on-ramps, diverging flows etc. This was later extended to general junctions. Literature references include (Smits, Bliemer, Pel, & Arem, 2015) and (Tampère, Corthout, Cattrysse, & Immers, 2011).

Capacity and storage constrained models form the most advanced of the spatial models, as it explicitly models spill back in the form of horizontal queues. This means that the storage space of the links is considered. Some examples of dynamic models that are based on this spatial consideration include: (van der Gun, Pel, & van Arem, 2017), (Raadsen, Bliemer, & Bell, 2016), (Gentile G. , 2010), (Daganzo C. , 1993), (Yperman I. , 2007) etc.

2.2.7. Simulation Based Macroscopic Models Used in Practice

Initially, most macroscopic models were used in the static setting. The major limitation of static assignment is its inability to fully capture the true dynamics of trip departure and real-time routing behavior with the risks of underestimating the congestion level. In addition, static assignment may result in link volume that exceeds link capacity. Despite these limitations, static assignment is an extremely valuable approach to traffic analysis as it allows to quickly estimate the use of traffic networks and to develop an initial appreciation of the situation. This initial best estimate can then be used to carry out more detailed analysis as well as the more demanding dynamic traffic assignment (OmniTRANS Transport Planning Software, 2016).

In the Netherlands OmniTRANS, is the leading static macroscopic modeling package, and is often used for municipal and regional traffic analyses. The OmniTRANS framework is developed to allow easy multi-class and multi-modal traffic modeling, which are among its main strengths in comparison to other macroscopic models. The usability of Ruby programming language gives the software an added edge for creating user-specific job scripts and solutions. The software has ventured into dynamic modeling through the StreamLine dynamic traffic assignment framework. Furthermore, the software includes a wide range of functionalities that allow for fast and in-depth analysis of results. Figure 4 illustrates the network loads in the Delft network, Netherlands. Being a dynamic model, the visualization can be obtained per time step of propagation as shown in the box in the bottom-left.

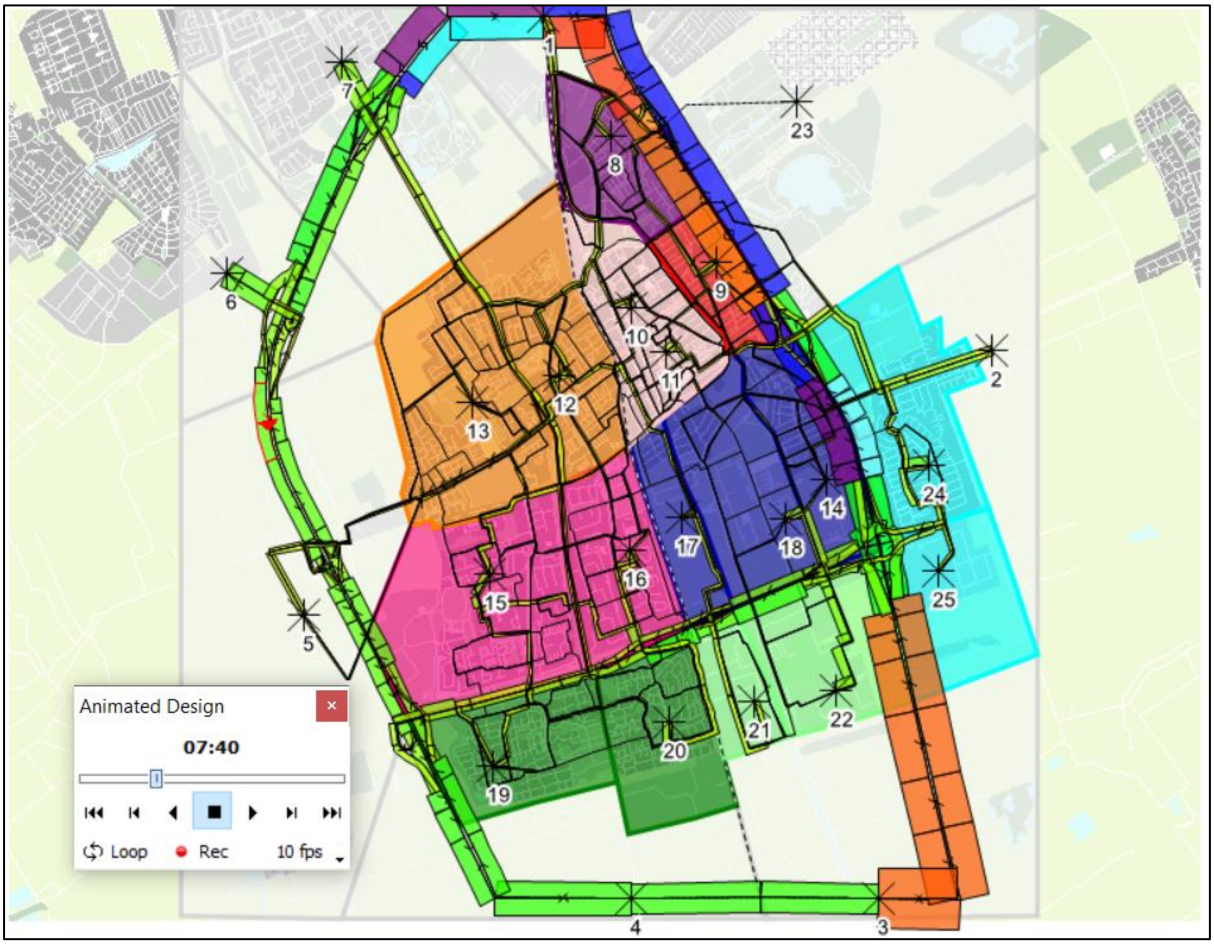


Figure 4: The macroscopic representation of Traffic in Delft Network, Netherlands. Source: OmniTRANS

Internationally AIMSUN is probably the best known dynamic macroscopic model. PTV VISUM (Germany), Cube Voyager (USA), EMME (CANADA), TransCAD (USA) and TransModeler (USA) are also among the leading model packages. INDY is another dynamic macroscopic model developed by TNO, Netherlands and KU Leuven, Belgium and is applied to various studies (Bliemer, Versteegt, & Castenmiller, 2004). MARPLE by Dr. Henk Taale, (Taale H. , 2008) is another Macroscopic DTA model developed to specifically target studies involving traffic management and anticipatory signal control. Detailed characteristics of this model will be described in the Section 2.3.1.

Most of these models make use of the LTM (Yperman I. , 2007) for propagation of the traffic flows. As previously mentioned, the LTM removes the space discretization errors posed by the CTM and has therefore gained popularity amongst model developers for application. However there exists other (Dynamic Network Loading) DNL mechanisms such as flow – travel time relations used in MARPLE and the second-order CTM used in StreamLine-MaDAM. eGLTM is an improvement developed for the LTM algorithm proposed by (Raadsen, Bliemer, & Bell, 2016). A StreamLine implementation of this algorithm has been developed and implemented in OmniTRANS software. Detailed characteristics of the StreamLine models will be described in Section 2.3.2.

A software package which is widely used for Macroscopic DTA applications in Europe and globally, is the PTV VISUM TRE, developed by the PTV group, Karlsruhe, Germany. It computes dynamic assignment where path choices and demand loading depend on the travel times obtained through any network simulation model which assumes the resulting node splitting rates (Calvert, Minderhoud, Taale, Wilmink, & Knoop, 2016). The propagation of traffic flows is based on an extension of LTM model, the GLTM – General Link Transmission Model. The implementation of the software is completely parallelized to maximize performance by fully exploiting the available computing resources.

The latest trends observed in model development is the possibility of combining attributes of different models into one. Such a fusion results in what is termed as Combined or Hybrid Model. Some combinations of modeling techniques are: microscopic junction modeling in macroscopic models, variable time intervals, integrated demand and assignment modeling etc. Aimsun offers the concept of Hybrid models through its mobility platform – Aimsun Next. It aims to combine microsimulation (including a pedestrian simulator), mesoscopic simulation, macroscopic functionalities, travel demand modeling and even two hybrid simulators (macro-meso and micro-meso) – all within a single software application (Aimsun, 2021).

2.3. Macroscopic DTA Models Under the Lens

The focus of this section will be about the three Macroscopic DTA’s that will be compared using the evaluation framework, namely: 1) **MARPLE**, 2) **StreamLine: MaDAM** and 3) **StreamLine: eGLTM**. The DTAs varies from each other in terms of propagation of traffic, algorithms used for traffic assignment to routes, treatment of junctions, size of time step used for network loading etc., which broadly covers the variability offered through macroscopic DTA models. The models chosen for the evaluation can be categorized according to the various classification schemes obtained from literature as stated in the preceding sections. Each DTA is sorted according one or more of these categories as further described. As a preliminary step in evaluation, model features and properties will be compared side-by-side as understood from the literature along with the algorithms applied within each model. Detailed studies on model performances will be analyzed while scoring the MoPs.

2.3.1. MARPLE – Model for Assignment and Regional Policy Evaluation

MARPLE is a route-based macroscopic DTA model used for integrated traffic management used at both local and global levels (Taale, Westerman, Stoelhorst, & van Amelsfort, 2004). The network scalability of MARPLE involves both that of highway and urban networks. The model was developed by Dr. Ir. Henk Taale as part of his PhD research. His research involved the interaction between traffic management measures and route choice in a traffic network (Witteveen+Bos & Taale, 2020). The target area of the model application was to regulate traffic by taking into account the route choice of the travelers in the form of anticipatory signal control. The categorization of MARPLE is different from the classification schemes mentioned in the preceding chapters. This is essentially because, the network loading in MARPLE is not based on a fundamental diagram (FD). However, it falls under the category of “*capacity and storage constrained*”, with model characteristics similar to that of a first-order FD based model.

MARPLE is essentially a model which simulates the supply modeling and the interaction between demand and supply. Thus, it uses demand input such as OD matrices, traffic networks and a list of parameters that is used to run the simulations. A detailed list of parameters and their uses are provided in (Witteveen+Bos & Taale, 2020). The basic structure of the model is shown in Figure 5. The model considers three components, the route set generation model, the dynamic route choice model and the dynamic network loading model.

2.3.1.1. Route Set Generation Model

The route set generation model determines a set of routes for a given OD pair with the transport network as an input. The basic assumption involved here is that the road user chooses a route from an a-priori-set of routes (in line with the route choice from a behavioral perspective). Furthermore, as the model is essentially route-based and uses route sets that are a priori generated, the computation time is greatly improved by avoiding time-consuming shortest path calculations (Bliemer & Taale, 2006). However, the disadvantage of the model is the fixed set of routes used throughout the model, which may not be realistic. Thus, a sufficient set of routes need to be generated for each OD pair.

Due to a priori generation of route sets, the number of unused routes must be kept minimum to avoid unwanted computation steps. The route generation model in MARPLE adapts a form of the "most probable route" approach provided in (Bliemer M. , 2001). The approach uses Monte Carlo simulations to generate routes in which the link travel times are assumed to be a random variable. Such an approach is assumed to model a scenario representing congestion. This method ensures a comparatively accurate route set, with lesser unrealistic routes. The assumption of free-flow travel time for routes can be avoided, and the computation time is relatively low even for large networks.

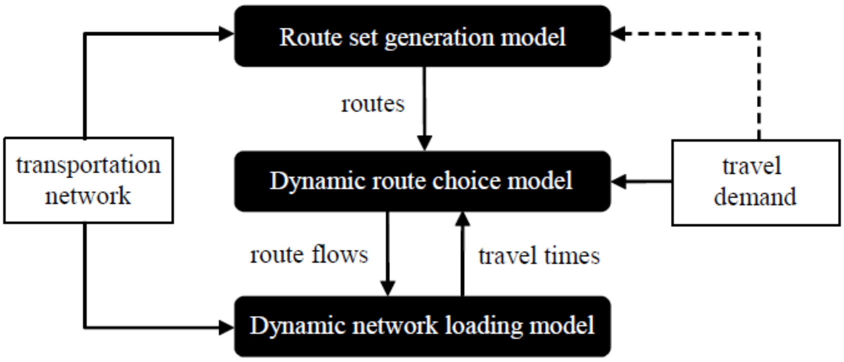


Figure 5: DTA components of MARPLE adapted from (Bliemer & Taale, 2006)

In route set generation, the link travel times τ_a for each link a is assumed be random variables:

$$\tau_a = \tau_a^0(1 + |\varepsilon_a|), \qquad where\ \varepsilon_a \sim N(0, \sigma^2). \tag{6}$$

where, τ_a^0 is the free-flow travel time and the random component ε_a is normally distributed with zero mean and variance σ^2. As per (6), a higher variance will lead to larger value of link travel times. Essentially if the standard deviation is assumed to be equal to 1/3, then the link travel times are never greater than twice the free-flow travel time (since Pr ($|\varepsilon_a|$) < 3σ = 0.997). Through iterations, the error term draws random values within the distributions to generate different route sets. Initially, the value of variance is set to zero, to ensure that the shortest path is always included.

MARPLE ensures that the detour is not too large by keeping the variance constant, resulting in only allowing routes that are within a certain threshold from the fastest route. An overlap filter is also used to remove routes which has too much overlap with the previously generated routes. MARPLE offers further input variables to reduce the number of routes generated. It is important to have a check on the OD pair demand where the route generation builds on. It is illogical to have routes for OD pairs having zero demand and hence it can be removed. Routes thus generated serve as an input for the dynamic route choice model.

2.3.1.2. Dynamic Route Choice Model

In the Dynamic route choice model, the route proportions and route costs are computed. Essentially this step describes the traffic assignment in the network. MARPLE incorporates three different approaches to network assignment as per classification scheme proposed in (Chen H.-K. , 1999) – A Deterministic Dynamic User Optimal, A Stochastic Dynamic User Optimal and a System Optimal Assignment. Interested readers are referred to (Taale H. , 2008) for a detailed explanation of each of these assignment techniques in conjunction with MARPLE.

The state of equilibrium in traffic modeling is an abstraction from the principle stated by (Wardrop, 1952), or in other words Wardrop's equilibrium. The First principle states that under equilibrium conditions traffic arranges itself in congested networks such that all used routes between an O–D pair have equal and minimum costs while all unused routes have greater or equal costs. This is under the assumption that all travelers perceive the same minimum cost and seek the same objective. In case of stochastic effects, the minimum travel cost is replaced by (perceived) minimum travel cost (Ortúzar & Willumsen, 2011).

For this research and in the rest of the report, we will assume that the route choice behavior is based on Stochastic Dynamic User equilibrium (SDUE). As commonly assumed, travelers will choose the route alternative with the (perceived) minimum travel cost (Bliemer & Taale, 2006). The reason for this assumption is that it is more realistic, compared to a Deterministic Dynamic User Equilibrium (DDUE). The realistic angle provided by SDUE comes from the usage of all routes by at least a few travelers as an outcome of stochasticity (a random term in route cost). Essentially the convergence for a DDUE is difficult to achieve, especially in case of a larger network. This is because the (perceived) shortest travel time used (for SDUE) updated at the beginning of each iteration are closer to the simulated route cost, which maybe far away from the shortest travel time (used for DDUE), while calculating the duality gap. The closeness to duality gap in case of SDUE is achieved due to larger demand spreading caused by the logit/probit choice model. Essentially, to achieve the conditions of Wardrop’s equilibrium, the duality gap values for DDUE should be zero. This makes convergence infeasible for DDUE, which motivates further, the choice of using SDUE during the simulation run.

The generalized cost in an SDUE may consist of route travel times travel, route toll costs, etc. For this thesis, we will be using only the route travel time. Each route cost may be perceived differently by different travelers. Hence the route cost is assumed to be a random variable by adding an unobserved random term. Under the assumption that all routes are independent and that the unobserved error term is following the Extreme Value Type I (or Gumbel) distribution, the route choice proportions are given by the well-known multinomial logit (MNL) model. The MNL models have been derived based on the assumptions that the error terms of the utility functions are Independent and Identically Distributed (I.I.D.). The I.I.D. property means that the sources of errors contributing to the disturbances must do so in a way such that the total disturbances are independent. In other words, the alternatives should not share unobserved characteristics (Ortúzar & Willumsen, 2011).

To tackle the problem of overlapping routes, a commonality factor (Cascetta, Nuzzolo, Russo, & Vitetta, 1996) is added to each route cost in MARPLE, denoted by F_p^{rs} (C-Logit Model). Then, the route choice proportions for OD Pairs *(r,s),* The route sets P^{rs} and cost of each route c_p^{rs} for each departure time *k* can be computed as:

$$\psi_p^{rs}(k) = \frac{\exp\left[-\mu\left(c_p^{rs}(k) + F_p^{rs}\right)\right]}{\sum_{p'} \exp\left[-\mu\left(c_{p'}^{rs}(k) + F_{p'}^{rs}\right)\right]} \tag{7}$$

Let the travel demand for OD pair (r,s) departing at time k be given by $D^{rs}(k)$. The route flows $f_p^{rs}(k)$ can be determined by:

$$f_p^{rs}(k) = \psi_p^{rs}(k)\, D^{rs}(k) \tag{8}$$

Route choice is an iterative process where the route flow variations occur as a consequence of changes in route choice proportions as computed through equations (7) and (8). Each iteration averages out the route flow values from the previous iterations through various averaging schemes. The commonly used averaging scheme in literature is MSA (Method of Successive Averages), where the weight used for averaging for every iteration step n, is given by $w^n = 1/n$. For MARPLE, this weight is calculated as $w^n = \alpha_1 \exp(-\alpha_2 n) + \alpha_3/n$. This modified averaging unit (MSA adjusted), ensures that a larger weight is allotted to a smaller value of n and a smaller weight for a larger value of n, thereby resulting in faster convergence (Bliemer & Taale, 2006). The values of constants α_1, α_2 and α_3 are 0.95, 0.25 and 0.05 respectively.

The convergence is usually checked using multiple criteria. Most models rely on absolute difference in route flows between iterations. However, this is not a metric that measures the proximity to UE conditions, as it only captures changes over iterations. Given that the averaging scheme (MSA or MSA adjusted) will reduce the change by definition (step sizes become smaller as the number of iterations increases) it is not a good metric for convergence. (Bliemer & Taale, 2006). A better measure for convergence is the dynamic relative duality gap as defined by:

$$G(k) = \frac{\sum_{(r,s)} \sum_{p \epsilon P^r} \sum_k f_p^{rs}(k)(c_p^{rs}(k) - \pi^{rs}(k)}{\sum_{(r,s)} \sum_k D^{rs}(k)\, \pi^{rs}(k)}, where\ \pi^{rs}(k) = \min_{p \epsilon P^{rs}} \left(c_p^{rs}(k) \right) \tag{9}$$

For each departure time k, this relative gap should decrease. However, in a SDUE, the value will never reach zero, but the iteration will stabilize to a value close to zero. The convergence criteria in such a case will be an input threshold for this value termed as duality gap. Due to this arbitrary condition imposed on the threshold, MARPLE looks at the change in route flows for every OD pair from one iteration to the next one. This change is also normalized using the OD demand. The maximum change over the OD pairs and time periods is compared with a threshold value. In practice this threshold varies between 0.1%-5% (Taale H. , 2008).

$$\max_{p} \max_{od} \max_{p \epsilon P^{rs}} \frac{\left| f_p^{rs}(k) - f_p^{rs}(k-1) \right|}{D^{rs}(k)} < \varepsilon^* \tag{10}$$

where, $\varepsilon^*$is the convergence error threshold. In practice this threshold varies between 0.1%-5% (Taale H. , 2008).

2.3.1.3. Dynamic Network Loading

The route flows are further assigned to the network using the Dynamic Network loading (DNL) model. It is the most computationally complex sub-model in the framework as it performs the task of propagating the traffic. The output of the DNL model is the link travel times and thereby the link costs. In MARPLE, the link travel time and delays are calculated based on link performance functions (travel time functions) (Taale H. , 2008). For the calculation of travel times in the links, four link types are defined: (i) normal links, (ii) controlled links - signal control or ramp metering, (iii) roundabout links, (iv) priority links (links ending at priority junctions). The travel time functions used for the links are derived from standard travel time functions. They depend on the junction type the link connects to (standard junction, signalized junction, roundabout or priority junction). The travel time function used for Normal links is derived from the Akçelik function (Akçelik R. , 1991) and (Akçelik R. , 2003). For controlled links, the HCM2000 (which is based on the work of Akçelik) function is used which is derived from (TRB, 2000). In the case of both roundabout and priority links, the calculation of travel times is two-fold: the determination of the capacity of the link and, based on that capacity, the calculation of the travel time. The travel time function used for both the links types is derived from (Troutbeck & Brilon, 2002). However, the capacity calculation defers between roundabout and priority links. Readers interested in detailed explanation of each link type are referred to (Taale H. , 2008).

At decision nodes, the traffic is distributed among the outgoing links according to the splitting rates determined by the route flows. Before traffic enters a link, the capacity check of the inflow link is conducted. If the demand at the node is greater than the capacity, the balance vehicles are stored in the upstream links. Such a redistribution can be termed as blocking back. Blocking back is modelled through the concept of 'available space'. The available space on a link determines how much traffic can enter the link and thus how much traffic is held back on the upstream links. This blocked traffic is distributed among the upstream links according to the number of lanes in MARPLE.

The splitting rates are calculated for every node, per time step and for every link-link combination, thereby covering the complete distribution over all the links and nodes throughout the propagation period. The splitting rate calculation is essentially a fraction between route flows. The numerator considers route flows over two consecutive links, which belong to the same rate. The denominator is all the route flows for applicable for the first link in the sequence, ensuring proportional distribution mathematically. The calculation is slightly different for origin links. Interested readers are referred to (Taale H. , 2008) for detailed mathematical formulae associated with splitting rate calculation along with an example to illustrate the calculation in a theoretical test network.

Traffic propagation in MARPLE takes place dynamically and it is calculated at link level, which is translated to route level using a trajectory method (Taale H. , 2008). This means that for every time step, several traffic variables are calculated in a specific order such as degree of saturation for the links, link delay and travel time, link outflow, node inflow, available link space, link inflow, node outflow, corrected link outflow, queue length and link flow for the next time step.

A special feature of the model is its ability to handle short links. Short links can be defined as those links, whose travel time to traverse with the input free flow speed is less than the time step used for model propagation. In other words, if the time is taken to go from the start of a link to its end, is lesser than the time step with which the vehicles are loaded in the network; within a single time step, the vehicles may occur at two successive links, which becomes an anomaly in traffic modeling. In normal cases, the time step of propagation is chosen such that its lesser than the time required to traverse such short links. However, in large networks, this can lead to slower computation. Another method to handle short links is to increase the size of the links. But this method comes with other problems such as increased vehicle kilometers, increased congestion and more delay than expected (Taale H. , 2008). To minimize these associated issues, MARPLE identifies such critical links (short links) at every time step and virtually lengthens it for propagation purposes. The virtual lengthening comes with adjustments to the link outflow from such critical links and the handling of congestion upstream. The lengthened link, which now has higher capacity, takes longer time to fill as opposed to the actual queuing formed in these links. Interested readers are referred to (Taale H. , 2008) for detailed explanation of these adjustment and associated test results for validating the adjustment.

For each of the nodes in the network, the inflows and outflows are computed based on route flows and splitting rates. Corrections for inflow due to capacity and outflow due to downstream queues are computed. The link flows and queues are readjusted based on this correction. The network loading in

MARPLE is link based. The route cost from the DNL model is fed back into the route choice model. In MARPLE, such an update occurs per iteration.

2.3.2. The StreamLine Framework – OmniTRANS

As mentioned before, OmniTRANS is an integrated package for multimodal and multi-temporal transport planning. In the Netherlands, the largest application domain for the software is for strategic planning. Developments on OmniTRANS started in 1997 by Dutch traffic consultancy Goudappel Coffeng. In 2003 a separate firm, OmniTRANS International (currently named DAT Mobility B.V.), was founded. The software development and marketing of OmniTRANS is performed by DAT Mobility.

OmnniTRANS offers a host of features with detailed modeling capabilities both for the demand and the supply side of planning. The focus of this section would be to touch upon details relating to the StreamLine framework, used by OmniTRANS for incorporating Macroscopic propagation models within its interface. The StreamLine framework aims to break the tradition by treating the DTA sub-models as individual building blocks instead of implementing the entire DTA framework as a single implementation together with network loading. This allows for re-use, modification and addition of the different sub-modules instead of re-implementing them with every network loading model/route choice model / averaging scheme / junction model etc. (Raadsen, Mein, Schilpzand, & Brandt, 2010).

The StreamLine framework is composed of the following elements as referred from (OmniTRANS Transport Planning Software, 2016):

- Input specification - The input is split into three parts: the transport network, the travel demand and a route set.

- A junction model for the detailed representation of both unsignalized and signalized junctions.

- The propagation model for dynamically loading the network and for modeling the dynamic propagation of the traffic across the network. Either use Madam or STAQ.

2.3.2.1. Route Set Generation Model

The focus of this sub-section will be to describe the route set definition for the model run. Similar to MARPLE, StreamLine offers a choice for the model user to provide an input file for the route sets or to allow the framework to generate routes. By default, route generation is based on the basic Dijkstra shortest path algorithm. StreamLine offers several options to include stochasticity in route set generation by invoking a Monte Carlo simulation (repeated random sampling) to generate alternatives for routes. Furthermore, the model also incorporates features to activate a set of adjustable filtering criteria that filter routes based on overlap in the route set and detours.

2.3.2.2. Dynamic Route Choice Model

As earlier mentioned, the route choice module provides information on how the traffic is distributed among various routes between a specific OD-pair. In MARPLE, this distribution is performed at the start of each iteration. However, StreamLine propagation models offer the model user to provide route choice moments in time during which the route fractions are calculated.

There are two route choice types available within the framework: ONESHOT- which defines the route choice for a single iteration run and MSA – which defines the route choice for multiple iterations and uses the Method of Successive averages for averaging over the iterations. Furthermore, the computation of route choice can be based on any of the below-mentioned options:

- **AON** – It is the most basic of the route choice, where the shortest route between a specific OD pair is allocated 100% of the route flows and the other routes remain empty.

- **Uniform Distribution**– A special type of route distribution, where the flows are spread uniformly amongst the available routes between an OD pair.

- **MNL** – A method to distribute traffic among routes in lines with the concept of SDUE as mentioned in the previous section. In this method the fraction of each available route is calculated based on the cost (travel time or distance) with a logit formula. The Multi-Nominal Logit (MNL) ensures the modeling of real-life situation, where not every traveler is aware completely about the shortest path between his/her origin-destination before commencement of the trip.

- **PCL** – A special form of MNL, but taking into consideration the overlapping of routes and not treating them as independent entities. Usage of this method is expected to obtain results closer to reality, by mimicking the effect of route independence. Readers interested in a comparison between PCL and C-Logit (route choice model in MARPLE) are referred to (Pravinvongvuth & Chen, 2005).

- **Input from Dataset** – StreamLine also allows the storage and extraction of route fractions in its data set, as input by the model user. Such extraction is usually used when the same traffic network is re-used.

The convergence in StreamLine is based on dynamic relative duality gap as provided in Equation (9), Section 2.3.1.2. In StreamLine, routes and route costs are computed for every mode or modelled user class accounting for junction delays, link cost functions and known conditions at the time of departure. Based on time-varying link speeds, link densities, route travel times and route travel costs are calculated. StreamLine supports both reactive (instantaneous) or predictive (trajectory based) travel times. The cost of travelling on a route is the summation of all the costs incurred while travelling on the links and turns composing the route. In a dynamic assignment, the cost of travelling on a route varies over time as a result of the varying traffic flows.

StreamLine uses a generalized cost function to calculate the cost of travelling on a link or turn at any particular time. The generalized cost has three components 1) Travel time, 2) Distance and 3) Additional static cost (links only), such as tolls. Similar to MARPLE, the default setting is to use the travel time. However, model user can choose to use the other components and custom defined travel impedances, while computing the generalized cost.

2.3.2.3. The Propagation Models in StreamLine

The propagation or DNL part of Streamline features three different models, as listed and briefed below. StreamLine models ensure that the capacity and storage constraints are satisfied, due to which they can be categorized as "*Capacity and Storage constrained*" classification as mentioned in the preceding chapters. Further on, the focus of the section will be on MaDAM and eGLTM as they align with the research scope:

1. StreamLine offers a macroscopic DTA model capable of incorporating all the features as described in the above section to match the results close to realistic situations. The DNL is composed of the MaDAM propagation model and the XSTREAM junction model. These

models are based on computationally efficient algorithms making the dynamic modeling of traffic flows in medium and large-scale urban networks possible within a reasonable amount of time.

2. A macroscopic semi-dynamic network loading model that uses a static demand and assumes traffic to travel instantaneously from origin to destination. This model is composed of the STAQ (STatic Assignment with Queuing) propagation model and is a two-phase model. The first phase, the squeezing phase, the traffic is assigned to the path keeping capacity in consideration. At every bottleneck traffic is queued vertically. In the second phase, the queuing phase, the traffic in the vertical queues is propagated upstream to represent shock waves (Brederode, Bliemer, & Wismans, 2010).

3. StreamLine has a third propagation model in its arsenal - the eGLTM in conjunction with the OmniTRANS junction capable of modeling traffic networks both at global and local levels. This event-based LTM proposed by (Raadsen, Bliemer, & Bell, 2016) does not rely on time or space discretization as is the case with MaDAM. This third model is relatively new to OmniTRANS and the full-scale implementation requires further testing and benchmarking, which is partly performed through this thesis project.

StreamLine: MaDAM

The propagation model - MaDAM is a second order, largely based on the METANET as described in 2.2.4. However, unlike the METANET, MaDAM uses the fundamental diagram based on car-following model proposed by (Van Aerde, 1995). The METANET model is designed for motorway networks hence it incorporates only merge and diverge nodes. In MaDAM, this is integrated with macroscopic urban DTA modeling. Such an integration uses a different anticipation term proposed by (Raadsen, Mein, Schilpzand, & Brandt, 2010). The anticipation term ensures the faster changes of speeds modelled to mimic the more aggressive urban driver compared to a vehicle in the motorway.

StreamLine: eGLTM

As previously mentioned, an important weakness of a CTM is spatial averaging errors encountered due to discretization of space. The Link Transmission Model proposed by (Yperman I. , 2007), eliminates spatial averaging errors by looking at link boundaries. The LTM uses Newell’s triangular fundamental diagram in its basic formulation and applies it in a network context. The Generalized Link Transmission Model (GLTM) proposed by (Gentile G. , 2011), generalizes the Yperman model to any concave fundamental diagram. However, the problems associated with discretization of time still existed. It may be noted that both LTM and GLTM models employ time discretization, hence temporal averaging errors occur in both models.

The event-based Link transmission model (eLTM) proposed by (Raadsen, Bliemer, & Bell, 2016) does not rely on time discretization and hence removes temporal averaging errors. In (Bliemer & Raadsen, 2019), the event based LTM is extended to a generalized continuous-time LTM formulation considering a FD with smooth non-linear branches. This led to the creation of event-based Generalized Link transmission model (eGLTM), which is a First order model, which also includes non-linear FD as opposed to a triangular FD used in eLTM.

Currently, the StreamLine implementation of eGLTM supports triangular (Newell) and Quadratic-Linear (QL) fundamental diagrams. The eGLTM algorithm, tracks changes in traffic states over space and time. The number of calculations is restricted only to flow rate changes on link boundaries and is based on cumulative vehicle flows. The innovation in the algorithm is the computation efficiency. The algorithm is capable of constructing cumulative inflow and outflow curves by exactly tracking the moments the flow rate changes. Detection of the change in flow rate is based on an input threshold. This helps to improve the efficiency of the algorithm by eliminating minor flow rate changes (Bliemer & Raadsen, 2019).

Readers interested in the mathematical formulation and analytical validations of the algorithm are referred to (Bliemer & Raadsen, 2019) and (Raadsen, Bliemer, & Bell, 2016). The StreamLine application of the eGLTM algorithm for propagation aims to yield more accurate results. As a research objective for the current thesis, the comparison between StreamLine: MaDAM and StreamLine: eGLTM will be interesting as the only difference between the two is the algorithm employed for propagation and junction modeling.

2.3.2.4. Node Model

For an LTM, apart from the link model, an essential part of a first order or second order DNL is the node model. StreamLine eGLTM uses the node model proposed by (Tampère, Corthout, Cattrysse, & Immers, 2011). Node models have two important functions in a DNL model - The first is to impose constraints on the outflow of each incoming link (limited supply of the node itself or node supply constraints); the second to seek consistency between the demand and supply constraints imposed by the incoming and outgoing links (and the node itself) (Tampère, Corthout, Cattrysse, & Immers, 2011). To ensure consistency between demand and supply constraints, a distribution of the available downstream supply over the incoming links has to be determined. The various constraints interact with each other and with the flows transferred over the node, which is captured by the Supply Constraint Interaction Rule (SCIR). This rule should represent the aggregate driver behavior at a congested junction. Node model instances for specific junctions are obtained by introducing a SCIR and node supply constraints. Together they ensure that the model captures realistic traffic flow over junctions.

On the basis of literature study, Tampère has identified a list of requirements which needs to be satisfied by any first order macroscopic node model in order to yield a realistic, consistent solution. The requirements are listed below for merge models and diverge models separately (FakhraeiRoudsari, Huang, & Tampère, 2015):

Merge

1. All flows should be positive at all times.

2. Continuity needs to be respected over the merge: the outflow should at all times be equal to the sum of the inflows.

3. outflow should never exceed the capacity of the receiving link, nor the max receiving flow of the receiving link (which can be lower than capacity eg in spillback conditions).

4. Neither of the inflows should exceed the capacity of the corresponding sending link

5. Neither of the inflows should exceed the sending flow of the sending link (that is, the amount of traffic that was able to reach the end of the sending link). This requirement essentially checks for FIFO (First-In-First-Out) rule.

6. There exists a Degree of Freedom (DoF) whenever the sum of inflows exceeds the receiving flow: there is then a constraint on the sum of the inflows, but ambiguity remains about the separate values of the terms in the sum.

7. This degree of freedom needs to be additionally specified, taking into account the following:

- *Realism*: the real reason to specify the DoF is that it best reflects real behavior at merging points. The macro assignment of outflow opportunities over the candidate incoming links is an aggregate of the underlying microscopic behavior. That is why for instance the number of lanes may be a good indicator of this assignment: if X outflow per time unit occurs towards a three-lane outgoing link, it is reasonable to believe that the sending link with two lanes will consume double of those flow compared to his single-lane sending link competitor in the merge.

- *Invariance principle*: the assignment of outflow over sending links should be invariant for substituting the sending flows of the incoming links with their corresponding capacities. Otherwise, the model would give unstable results (flip-flop). Eg If a rule based on ‘demand’ (=sending flow) will not be invariant: constrained receiving flow is assigned over the sending links according to the sending flows; suppose this causes a queue to grow on both sending links. Now they would both send capacity towards the merge (discharge from queuing), which would yield a different assignment of receiving flow over the sending links. Cases can be constructed in which this new assignment makes one of the incoming queues dissolve, after which the process repeats in oscillation.

- *Flow maximization*: each flow in the merge is limited by either the sending flow, or by the receiving flow (or exceptionally by both); no flow should be strictly lower than all of its constraints; in other words: all flows should be actively constrained by either merge demand (sending flow) or supply (receiving flow).

Diverge

1. Same positivity, continuity, capacity and flow maximization requirements as merge.

2. The ratios of the turning movements at nodes (called turning fractions) need to be respected in the solution of the model. Eg. if 1 out of 4 vehicles in the sending link want to turn left, and the left receiving link can only accommodate half of this turning flow, then also the outflow from the sending link towards the other outgoing links will be restricted to half of the sending flow. Note that this is the same as requiring FIFO (first-in first-out) at the link end, as if no overtaking is possible. The reason why FIFO should hold near the node is that otherwise FIFO could be violated on the OD-level: a vehicle leaving its origin later may overtake its predecessors at the non-FIFO node and arrive earlier in its destination; this conflicts with the typical equilibrium conditions, where a preceding vehicle can now arrive earlier at their destinations by leaving later, which yields an inconsistent result → model.

The node model proposed by (Tampère, Corthout, Cattrysse, & Immers, 2011), satisfies the above requirements. These requirements are further checked during the research for the different Macroscopic DTA’s.

2.3.2.5. Junction Modeling

StreamLine offers the feature to model junction delays, which leads to more realistic assignment results, especially in urban areas. This further permits the analysis of various junction measures at network level. The junction model in StreamLine: MaDAM – XStream, provides an additional layer of abstraction to the Propagation Model to facilitate the modeling of both unsignalized and signalized junction within a DTA model.

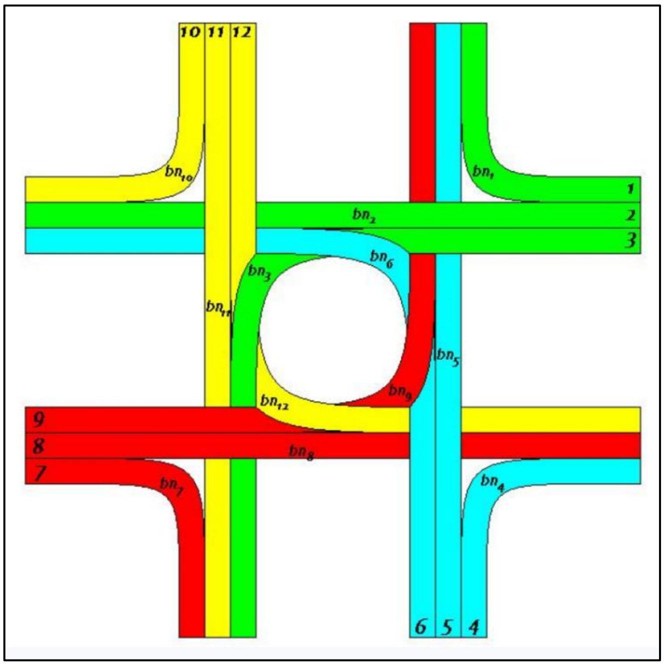


Figure 6: Possible turns in a four-way junction with bottlenecks. Source: (OmniTRANS Transport Planning Software, 2016)

XStream simulates every turn separately by reserving a segment of some length for this turn in the macroscopic DNL model. On this segment, both speed and capacity are adjusted according to the junction-type and the intensity of opposing traffic. XStream utilizes an adaptation of the static junction theory to determine the general delay and exit capacity on each turn for each simulation step.

There are three different unsignalized junctions supported by XStream: equal, give way and roundabouts. The formula for the calculation of the mean delay is divided into three parts: the uniform, incremental and geometric delay and the calculation is compliant with the current international standards stated by the Highway Capacity Manual 2000. A similar methodology is applied to model signalized junctions. In this case, the turning capacity is influenced by the green time and the cycle time.

Figure 6 depicts all possible turns of a four-way junction. Every turn *i* has some kind of bottleneck, bn_i which has a given length, capacity and maximum speed depending on the value of the conflicting flows and the specifics of the junction. The schematic view is always the same, whether the junction has traffic lights, is an all stop junction, roundabout or another type of junction. The only difference is the formula for the bottleneck bn_i. This extra layer of abstraction introduced in XStream is able to deal with all junctions in the same way while still being able to mimic the junction specifics defined by its bottlenecks.

2.3.3. Qualitative Comparison of the Models

From the above understanding it is clear that models selected for the evaluation in the current thesis differ in the model structure, the propagation algorithms and the features entailed for different transport planning applications. The differences in the modeling properties are noted down and summarized in Table 1.

This comparative table forms the primary step in evaluating the models. The comparison will further motivate the formation of various MoPs which will be used in the evaluation framework. Literature on motivation for forming the evaluation framework is provided in Section 2.4

Table 1: Summary of the difference between the three Macro DTA models

Components	**MARPLE**	**StreamLine: MaDAM**	**StreamLine: eGLTM**
Route Generation	Input route set or generation using Monte Carlo approach. Overlap filter for removing routes with route overlap.	Accelerated Monte Carlo approach - filtered based on route overlaps and detours	
Route choice models	C-Logit, AON	MNL, PCL, AON	
Propagation model	Link travel time & delay computation based on link travel time functions, further assigned at route level.	2^{nd} order CTM – extension of METANET	eGLTM
Fundamental diagram	*NA*	Van Aerde, Smulders, Newell	Newell, Quadratic-Linear
Explicit Node model	Uses the node model proposed by (Tampère, Corthout, Cattrysse, & Immers, 2011). However, Conservation of Turn Fractions is not strictly followed.	*NA*	Uses the node model proposed by (Tampère, Corthout, Cattrysse, & Immers, 2011)
Averaging schemes for iterations	One-Shot, modified MSA (refer to Section 2.3.1.2)	One-Shot, MSA	One-Shot, MSA
Junction modeling	Travel time functions based on Junction type - standard junction, signalized junction, roundabout or priority junction	StreamLine XStream Junction Model	OmniTRANS Junction Model
Traffic Management controls	➢ VMS (Variable Message Sign) ➢ Ramp Metering ➢ Lane Adapter (opening and closing of peak hour lanes) ➢ Anticipatory control scheme in traffic signaling for incorporating user choice.	➢ VMS (Variable Message Sign) ➢ Ramp Metering, ➢ Lane Adapter (opening and closing of peak hour lanes) ➢ Dynamic Link Attribute Adapter (Road work, Variable speed limits, weather changes) ➢ Outflow Limiter (Bridge Opening, train crossing)	(Under Development)

2.4. Evaluating Macroscopic DTA Models

This section provides a brief on various literature findings used for creating the framework. Based on the literature survey, there exists no framework or methodology to evaluate Macroscopic DTA models qualitatively. Numerous articles exist in the field of evaluating and validating the model results based on empirical data. Therefore, the literature review in this topic was done specifically to adopt some features of the evaluation framework (Section 2.4.1) and the motivations for incorporating the MoPs which act as yardsticks for evaluation (Section 2.4.2).

2.4.1. Literature references for Evaluation Methodology

(Ni, Leonard, Guin, & Williams, 2004) explains a model life cycle in their systematic approach to validating simulation models. Validation here would typically refer to checking the goodness of fit between the model simulated results and observed empirical results. As one of the initial steps towards validation, the authors emphasize the need to enlist the Measures of Performances or MoPs – which refers to the target variables on which the model assessment is based on. After this step, the authors refer to qualitative and quantitative techniques to compare model results. Qualitative techniques, also known as subjective, visual, or informal techniques on some other occasions, are typically performed based on visual comparison of the predicted and observed data in various graphs and plots. It is generally adopted as a preliminary technique to evaluate model performance and identify problems. Some of the qualitative methods described were – series plot, contour plot, surface plot, diagonal plot, histograms, animations etc. Detailed explanation of various quantitative validation techniques is described in (Ni, Leonard, Guin, & Williams, 2004), which will not be covered in this literature review.

(Rao, Owen, & Goldsman, 1998) argues that the purpose of the model determines what aspects of the model to validate and their levels of detail. In this regard, the paper presents a multi-stage validation framework which is primarily distinguished into operational validation and conceptual validation. The operational validation process involves comparisons between model predictions and measured real-world system behavior. Conceptual validation is a qualitative assessment of a model's theoretical underpinnings, as well as its implementation, evaluated in the light of sound and accepted theoretical methods. Conceptual validation is further distinguished into two methods – model survey and model walkthrough. Model survey as argued by (Rao, Owen, & Goldsman, 1998) introduces the requirement to engage the end user perspective in the analysis of model validation results and methodology. The authors describe this community of model users as three – Researchers, Developers and Practitioners. The survey described involves a questionnaire being sent to the members of this community for collecting the responses. In a model walkthrough, the logic and documentation of the model are reviewed and criticized constructively to suggest improvements or to re-check the steps involved in the model working.

(Flügel, Flötteröd, Kwong, & Steinsland, 2014) shows a tabular form of model evaluation, Macroscopic - Static and Macroscopic/Hybrid - Dynamic models are compared and evaluated under various MoPs for Strategic planning applications. A similar performance evaluation is carried out by (Salgado, Jolovic, Martin, & Aldrete, 2016) wherein three microscopic DTA models – Aimsun, VISSUM and TransModeler are compared qualitatively and are scored from 1 – 5 under various model/software properties such as Graphic User Interface (GUI), vehicle routing capabilities, driving behavior parameters, run time etc. The scoring is performed by the authors based on their experience in modeling Ysleta Zaragoza border crossing. The preliminary comparison here was conducted by noting down the strengths and weaknesses of each model. A performance evaluation matrix with scores was the final output of their research, with a recommendation of a specific model for a specific application.

A detailed extension of a matrix form of traffic model evaluation is showcased in (US Department of Transportation, 2004). The framework provides 7 criterions for selecting a traffic tool – Geographic scope, Facility type, Travel mode, Management strategy, Traveler response, Performance measures, tool/cost -effectiveness etc. Each of these criteria is checked for the different stages of transport projects such as planning, design and operation/construction. The next dimension in the framework is the analytical context in which the evaluation is applied, which in other words refers to the particular part of the model which is being analyzed such as travel demand module, traffic optimization module, macroscopic simulation etc. The scoring of the categories involves the use of weights for the context and final ranking based on weighted sum averaging. The final output of the framework is the analytical context module most suitable for the intended criteria and stage of the transport project.

2.4.2. Literature references for MoPs

Four literary sources were identified and closely studied to formulate the evaluation MoPs as described below:

1) (Bliemer M. , Raadsen, Smits, & Romph, 2013) describes the desired properties for traffic assignment models for strategic transport planning applications which can be broadly listed as below:

    a) **Realism of results**: The closeness of the results of the DTA to that of actual observed behavior is achieved through this property. The authors study the various aspects required to achieve realism in the main sub-modules of DTA – The dynamic route choice and traffic flow propagation.

    b) **Robustness of results**: The authors argue that it is of utmost importance for strategic transport models to have stable results within a specific scenario or variant i.e., the need for a model to be robust. A model is said to provide robust results if marginally different inputs only lead to marginally different outputs.

    c) **Consistency of results**: The need for consistency among models of various spatial levels of detail is emphasized through this property. The authors describe the need for mesoscopic models results as an aggregation of microscopic models. Similarly, macroscopic model results are expected to be consistent with the aggregated results of mesoscopic simulation. Microscopic model results are treated to be the baseline for comparison, as they are widely used for operational planning applications.

    d) **Reliability and accountability of results**: Refers to the ability with which the model results can be explained on the basis of the underlying mathematical principles and thereby prevent the model from becoming a black-box.

    e) **Ease of use**: Emphasize the importance of having model results within a feasible run time and computational complexity.

2) (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019) provides an extension to the abovementioned properties by considering the large-scale application of traffic assignment models for strategic planning applications. Within this extension, the authors established the following desired properties for the same:

    a) **Tractability**: The extent to which calculations in each model component can be verified using the theory behind the component or sub-model.

b) **Accuracy under congested conditions**: The extent to which flow metering, spillback and route choice effects caused by congestion are included in the model.

c) **Accountability**: The extent to which different model components can be isolated and verified

d) **Robustness (1)**: The extent to which the model is free from random variables that affect its outcomes.

e) **Robustness (2)**: The extent to which the model converges to a defined and meaningful stable state.

f) **Computational efficiency**: The extent to which run times and memory requirements are acceptable for calibration and application of large-scale models.

g) **Input requirements**: The extent to which input requirements are available with acceptable uncertainty for distant future scenarios.

h) **Applicability**: The extent to which the model is applicable for all vehicle classes and for both urban roads and motorways.

3) (Ortúzar & Willumsen, 2011) provides the list of requirements for truly dynamic traffic assignment model as described by (Heydecker & Addison, 2005):

   a) **Positivity:** DTA models are only interested in non-negative flows on links, paths, trip matrices and costs.

   b) **Conservation**: the model must satisfy flow conservation requirements.

   c) **First In, First Out (FIFO)**: in real-world single-class traffic network, the FIFO behavior generally prevails and this must be maintained in the model if proper delays are to be estimated.

   d) **Minimum travel time**: flows do not propagate instantaneously.

   e) **Finite clearing time**: there are no queues left at the end of the modeling period; infinite delays do not occur (as a standard queueing model might suggest).

   f) **Capacity**: there is such a thing as strict capacity constraint in the sense that actual flows cannot exceed it even for a short period of time.

   g) **Causality**: delays now are affected by what other vehicles do or have done in the past, not in the future.

4) (Chiu, et al., 2011) describes the defining quality of DTA model outputs as three dimensions mentioned below:

   a) **Convergence**: Almost all equilibrium-seeking DTA algorithms adjust the route assignment using an iterative solution procedure. Among the methods used for convergence, the authors argue “dynamic relative duality gap” to be intuitive and sound. As earlier mentioned, the solution is assumed to be converged, when the relative duality gap has reached a prespecified tolerance level.

b) **Solution Sensitivity and Stability**: The authors approach the question of difference in results between a base variant and alternative variant in a unique way - if the solution to an alternative variant exhibits unexpected features, this may indicate that a poor approximation to equilibrium has been computed for the base or the alternative, or both. They argue that for a scenario-based comparison to be valid, the individual equilibrium solutions must be computed to a precision that is greater than the differences between the solutions of the alternative problems; otherwise, any real differences between the alternatives will be lost in the imprecision of the calculated solutions.

c) **Realism of Traffic Dynamics**: DTA model outputs from a specific traffic network can easily be validated with empirical data from actual observed count, mainly due to its dynamics involved in space and time. This is considered to be one of the most prominent qualities of a DTA compared to static models.

2.5. Concluding Remarks

The focus of this literature overview was on the theoretical aspects of a DTA model and the MoPs related to this. Aspects that are not covered in the current research (due to scope definition) but still important are the data requirements, model functionality, transparency, etc. which are directly related to model users such as clients and policy makers. Furthermore, evaluation in the lines of different levels of validity such as verification, face validity, construct validity (ability to calibrate the model), and predictive validity are also important considerations while evaluating traffic models.

As stated before, the literature review performed for creating the evaluation framework was to mainly inspire the researcher to adopt certain aspects of other frameworks and methodology. Specific ideas from existing frameworks have been reshaped to suit the need of the current research objective. The detailed methodology of the framework along with the specific literature from which the inspiration was drawn is described in the next section.

3. Methodology – EMMa

The following section will describe in detail the methodology used in development, operation and application of <u>**EMMa** – Evaluation Model for Macroscopic DTAs</u>. From the literature review it was understood that there is requirement of a multi-dimensional framework for evaluating Macroscopic DTA models. This led to the creation of EMMa. Interestingly, "Emma" is an English name with roots in an old Germanic word meaning “wholistic” or “universal (Wikipedia, 2021). As the name suggests, the framework is developed by taking into consideration a wholistic perspective. The aspect about the framework which makes it multifaceted is its dimensions, which will be explained in detail in this section. The structure of EMMa is inspired

While explaining the dimensions, details on scoring, weights and the theoretical test networks used will be covered. MoPs related to scalability will be tested on a real-world network. The details of which will be provided in the results (See Section 4). Departing from the dimensions of EMMa, which forms the building block for the evaluation tool, the subsequent section will cover the working of EMMa including the standardization technique adopted for scoring the MoPs, which ultimately will provide the DTA model rankings. Throughout the methodology, the term “modeler” is used to define the person using EMMa to evaluate the different DTA’s.

3.1. Dimensions of EMMa

The multiple dimensions in EMMa adds to the elasticity for evaluation by taking into account various perspectives (Figure 7).

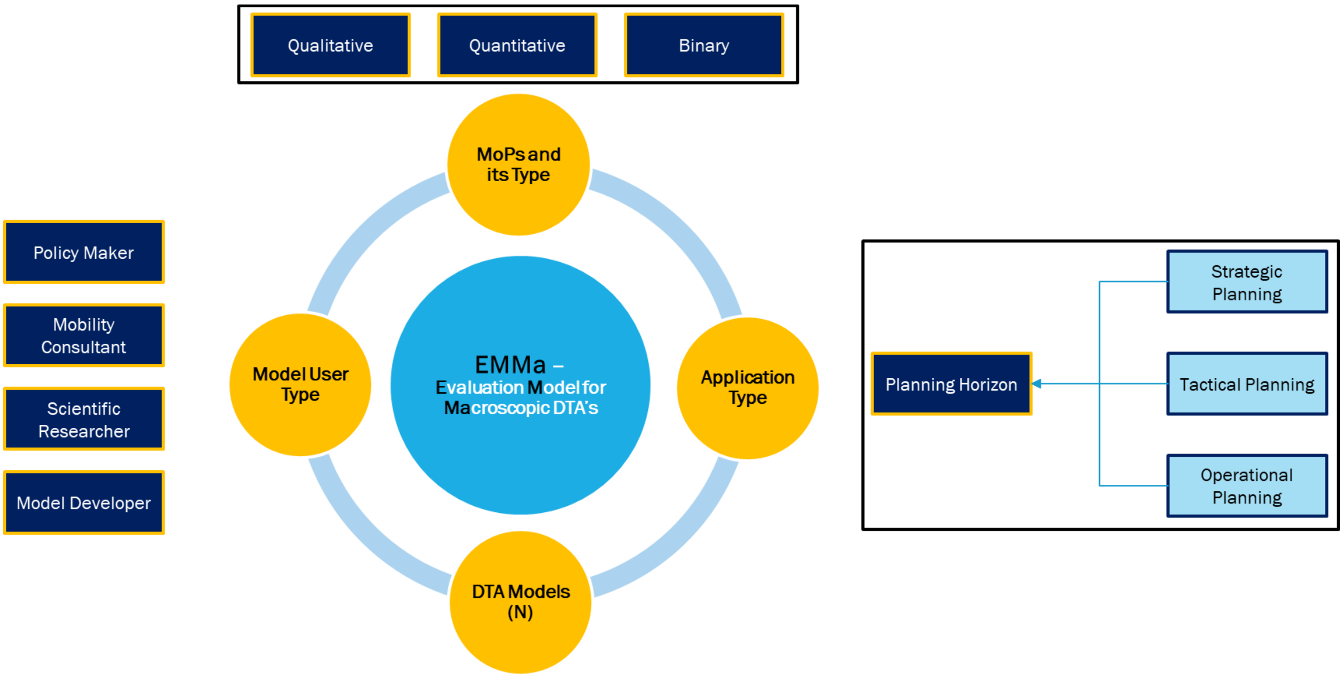


Figure 7: Dimensions of EMMa

EMMa is governed by four dimensions 1.) The MoPs and its type (Primary Dimension), 2.) The model user type (Secondary Dimension), 3.) The application planning horizon (Secondary Dimension), 4.) the DTA models in itself used for evaluation. (Primary Dimension).

The secondary dimensions in EMMa provides the model user with the flexibility to evaluate the models from various perspectives. The exact method for this sensitivity will be explained in the Section 3.5.

The following sections will describe in detail each of the dimensions, the inspiration to use the various sub dimensions within them for EMMa, literature references and criteria to use them. The DTA models involved for the evaluation require some prior literature study before application, as described in Section 2.3.

3.2. Measures of Performance (MoPs) and its Types

The MoPs form the core of the framework and acts as the primary dimension. 28 MoPs have been identified after performing reviews from different literature sources, considering the various aspects of a typical DTA Model. The large number of MoPs ensure that DTAs are evaluated at various aspects ranging from key modeling properties to ease of their application. Based on this motivation, they have been shaped and classified further to align with Macroscopic DTA modeling. Due to constraints of time and availability of data, the MoPs related to empirical validation have been removed from the present study, even though they form an essential part of evaluation as described in the preceding sections. The categorization of MoPs is performed in two layers:

- *Role of the MoP:* Describes the aspect of the DTA model which is being studied. The MoPs are further classified into 7 sub categories, and their definitions have been adapted from existing literature as listed and referred below:

    a. **Conceptual Validation**: The comparison of the effects observed in a real transport system to the simulated model. This involves checking the presence of real-world effects such as blocking back, capacity drop, smooth speed variations, stop and go waves etc. in the results of the simulated model (Ni, Leonard, Guin, & Williams, 2004). Due to the lack of empirical dataset, the comparison is performed based on the theoretical expectation.

    b. **Model robustness**: The extent to which the model converges to a defined and meaningful stable state (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019).

    c. **Applicability**: The extent to which the model is applicable for all vehicle classes and trip purposes. Adapted and modified from (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019).

    d. **Tractability**: The extent to which calculations in each model component can be verified using the theory behind the component or sub model (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019).

    e. **Integration of Network Hierarchies - Urban and Motorway roads**: The ability of the model to handle traffic propagation in urban and non-urban links. The MoP is based on the expectation of having speed fluctuations in the two link types as a direct consequence of different free-flow speeds (Raadsen, Mein, Schilpzand, & Brandt, 2010).

    f. **Computational efficiency**: The extent to which run times and memory requirements are acceptable for calibration and application of large-scale models (calibration omitted from current study) (Brederode, Pel, Wismans, de Romph, & Hoogendoorn, 2019).

    g. **Usability**: The ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component (definition quoted from - IEEE Std.610.12-1990) (Seffah, Donyaee, Kline, & Padda, 2006).

- *Type of measurement:* The classification is based on the method used in computing the MoP. As a result, 3 sub-categories are identified as stated below:

    a. **Quantitative Measurement**: The measurement of these MoPs forms a direct input in EMMa, by means of the actual value obtained by running tests in the traffic network (Theoretical or actual use case).

    b. **Qualitative Measurement**: Almost 70% of the MoPs in EMMa are measured qualitatively. The modeler is provided with a method to score these MoPs on the basis of existing literature studies or through test simulations, comparing the expected results based on literature and theoretical model working to the actual simulated results. Details on evaluation method are described in detail against the explanation of each MoP in the subsequent sections. Dependency on the expertise of the modeler is minimized, by making him/her experience the working of DTA and its sub-modules under study. The qualitative score system used for the evaluation is as provided in Table 2. The scoring ranges from 0-4, each value corresponding to the description provided in the table. This aggregate range is expected to prevent confusion and uncertainty for the modeler while providing the scores.

Table 2: Qualitative score system used in the EMMa

Score from	Score to	Description
0	1	*The property under consideration is absent from the model*
1	2	*Instances of the property can be observed. However, the results are inconsistent and non-reliable*
2	3	*The model results showcase the presence of the property under consideration partially.*
3	4	*Model results completely in-line with the model property under consideration*

    c. **Binary Measurement**: MoPs which are based on binary measurements provides information regarding the presence of the modeling property or feature of the DTA under consideration. As the name suggests, a value of 1, indicates that the feature is present and a value of 0, indicates that the feature is absent.

As described before, the classification of the MoPs is performed in two layers and the primary distinction is based on the role of MoPs which will be described and explained in detail in the following sections. Additionally, while describing each of the MoPs, the type of measurement will also be described, which forms the secondary classification.

3.2.1. Conceptual Validation

Even though the research excludes the MoPs related to empirical validation, it was essential to check if the DTA model under study was able to mimic and simulate effects that are observed in a real-world traffic system. However, we restrict this checking to results from model simulations and compare this with expected results. In this regard, 7 MoPs were identified as listed below:

- *Flow Metering (strict capacity constraint) - Qualitative*: Adopting from classification scheme provided in (Bliemer M. C., et al., 2017), the need to restrict the traffic flow in a link strictly up to the capacity is an important property to consider. Flow metering occurs when this restriction applies and excess traffic (demand - capacity), is blocked from entering the link. Flow metering by itself creates vertical queues in the network, in links upstream of the critical links (where demand exceeds capacity). Vertical queues can also be termed virtual queues, where the excess traffic is stored or stacked virtually. The MoP being qualitative in nature, scoring is performed after running tests on the network (theoretical or actual use case) and as per Table 4. Please note that throughout the research, moving bottlenecks are excluded as most Macroscopic DTA models are unable to simulate them. Therefore, we restrict the scope to stationary and temporary bottlenecks.

- *Traffic Spillback (strict storage constraint) – Qualitative*: The MoP has also been adopted from (Bliemer M. C., et al., 2017). In conjunction with a strict capacity constraint, a strict storage constraint in the model will ensure that the flow does not exceed the capacity by diverting traffic to alternative routes or links with spare capacity or by buffering vehicles in the form of queues to links upstream. This horizontal queue formation in links upstream from the critical link can be termed as traffic spillback or blocking-back. It is important to note the difference of this MoP to the previous one, as it involves the constraint of available space in a traffic link. Such a phenomenon is prevalent in most real-world traffic networks where congestion in a road stretch would eventually trigger traffic blockages in those roads connected to the congested road. The extent of spillback is scored and evaluated by comparing the simulated model results to the theoretical expectation, thereby help in validating the model conceptually.

- *Capacity drop – Qualitative*: When the congestion due to a bottleneck is resolved, drivers tend to maintain a larger distance with the predecessor vehicle, which would ultimately lead to a decrease/drop in lane saturation flow and consequently the road capacity, which occurs in a real-world case. Similar to the previous MoP, the extent of this drop in case of a macroscopic DTA model is evaluated and scored by comparing the simulated model results to the theoretical expectation. Therefore, the scoring of the MoP requires running simulations on test networks or actual use case networks.

- *Smoothness of temporal speed and flow variations- Quantitative*: The variation of speeds and flows over time can be studied to check the closeness of the driving behavior to a realistic setting. The logic here would be to assume that distributions which are smooth over time, depict a more realistic variation in traffic states. The performance indicator used for this purpose measures the roughness of the distribution and is widely used in the field of material sciences (Gadelmawla, Koura, Maksoud, Elewa, & Soliman, 2002). Of the list of indicators, we choose the arithmetic average height (Ra), as it is self-intuitive regarding its definition, application and measurement. It is defined as the absolute deviation of the roughness irregularities in a distribution from the mean line over one sampling length as shown in Figure 8.

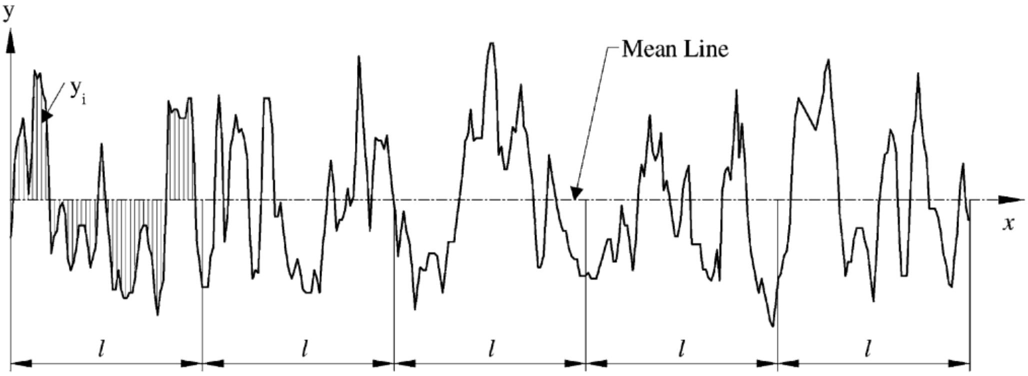


Figure 8: Definition of Arithmetic average height (Ra) adopted from (Gadelmawla, Koura, Maksoud, Elewa, & Soliman, 2002)

The formula for calculating the arithmetic average height is given by:

$$Ra = \frac{1}{N}\sum_{i=1}^{N}|z_i - m| \qquad (11)$$

Where N is the total number of traffic state values, i describes the time step of propagation for a specific link, z_i is the traffic state value at ith time step and m is the arithmetic average of all the traffic state value. From the definition and equation (11), it can be inferred that larger the roughness value of Ra, the distribution is less smooth. Two different Ra values are calculated in EMMa, for flow and speed distribution over a specific link. The choice of link to analyze will be decided by the modeler. In principle, a link upstream of a bottleneck will be expected to have variations in speed and flow throughout the simulation period and will be an ideal candidate link to choose from.

- *Presence of variable route set (Binary):* The option for the model to incorporate the feature of variable route options is checked and evaluated in a binary system.

- *Modeling of stop and go waves (Binary)*: In reality, it is possible for the front of the queue to keep moving, while the back of the queue moves backwards. This forms a series of high peaks and low valleys, observed in a space-time diagram (Taale H. , 2008). In these short traffic jams, vehicles come to (almost) a complete standstill. The duration of the queues is a few minutes. At the downstream end of the jam, there is no physical bottleneck, which makes these traffic jams be termed as "Phantom Traffic Jams". The possible reason for such a short traffic jams are small disturbances, like individual cars changing lanes or slowing down at a curve, which are absorbed by other drivers' adjustments. However, the presence of this effect needs to be checked in the DTA model results and is scored and evaluated in a binary system, similar to the previous MoP. Space-time diagrams of the simulated results are checked for candidate test scenario, as will be explained in Section 3.2.4.

3.2.2. Model Robustness

For the purpose of this research, Dynamic relative duality gap is used to find the convergence as defined and motivated in section 2.3.1.2. The relative gap is an estimate of the distance between the current solution and the optimal equilibrium solution. The primary assumption here would be that if the duality gap achieved is less than the input threshold by the modeler, the convergence has resulted in a stable state of equilibrium.

The question then becomes what should be an ideal value for this threshold. The general guideline is to make sure that the user benefits in terms of percentage time savings, are at least 10 times the relative gap (in %) (Ortúzar & Willumsen, 2011). However, (Boyce, Ralevic-Dekic, & Bar-Gera, 2004) investigated this guideline in some practical cases and recommended that the threshold be at most 0.01% (0.0001) for satisfactory convergence. This has proved to be an exacting requirement, creating abnormal model run time (Ortúzar & Willumsen, 2011). Although not applicable to DTA models specifically, (Patil, Ross, & Boyles, 2021) show that the threshold value should be dependent on the intended model usage (level of aggregation of output considered). The MoP used in EMMa is quantitative and the value input is the value of relative duality gap as mentioned in Equation (9), Section 2.3.1.2. After a fixed number of iterations, the model with relative gap value closest to zero will have a higher score.

However, for the purpose of comparing MAPRLE with the StreamLine models, a modification to the above-mentioned MoP is required. This is because in MARPLE, the relative difference in flows is used as a measure of convergence given by Equation (10), Section 2.3.1.2. To compare the results, the relative change in duality gap function is monitored in every iteration for the StreamLine models. The relative change in gap function is adapted from (Taale H. , 2008), as given below:

$$\frac{|G(k) - G(k-1)|}{G(k)} < \varepsilon^* \tag{12}$$

All three models are run for a fixed number of iterations on the same network, and the stabilization to an equilibrium is analyzed. The relative change in gap function (in StreamLine models) or change in route flows (in MARPLE), for the final iterations is used as a quantitative MoP in EMMa and is called convergence error (Taale H. , 2008). The smaller the value, the closer the model is to equilibrium. It may be noted such a measure of relative difference is valid only for a comparative analysis, as SDUE does not achieve true equilibrium (relative gap = 0).

3.2.3. Applicability

As discussed in the definition, the ability for the model to include features of multi-class (multi-modal or multiple user class) is evaluated through this category of MoPs. In this regard two MoPs are identified as listed below:

- *Difference in Network Supply based on Modes (Binary)*: The MoP checks the option in the model to include multiple modes within the network. The option usually entails the difference in free flow speeds, saturation flow of the lane, speed at capacity, number of lanes etc. which is usually specified for the link (network supply). If the option is present in the model, a score of 1 will be awarded and a score of 0, if otherwise.

- *Difference in Input parameters based on different user classes* (Binary): The option to have multiple user types (traveler type) with various trip purposes is checked through this MoP. Travelers exhibit a different value of time and value of distance depending on their trip purpose. In line with the previous MoP, if the option is present in the model, a score of 1 will be awarded and a score of 0, if otherwise.

These MoPs can also be scored qualitatively, by running simulations in test networks, which are mode-specific and user-specific, thereby comparing the model results with theoretical expectations. However, for the scope of this research, the evaluation is restricted to binary scores.

3.2.4. Tractability

This category of MoPs checks the proximity of the model results with the theory behind the calculations in each of the model sub-component. The method used to verify this proximity is through running simulations in test networks. In this regard, a total of 8 MoPs have been identified, which are scored qualitatively as per Table 2. The MoPs are as listed below:

- *Propagation - Link flows*: To evaluate the model performance for network loading module, during link propagation by comparing the expected theoretical results vs actual simulated results.

- *Propagation – Queuing*: To evaluate the model performance for network loading module, during queue formation by a stationary and temporary bottleneck, by comparing the expected theoretical results vs actual simulated results.

- *Propagation - Effect of link-level traffic controls*: To evaluate the model performance for network loading module, while varying some input properties of the fundamental diagram of a link such as speed, capacity (saturation flow) etc., by comparing the expected theoretical results vs actual simulated results. In other words, the effect of traffic control mechanisms such as dynamic link attribute editor, dynamic route information panels, lane controls, variable speed limit etc. is checked and evaluated.

- *Node model-merge & diverge behavior*: To check the compliance to requirements of a node model as stated in section 2.3.2.4 , by comparing the expected theoretical results vs actual simulated results.

- *Signalized Intersection*: To evaluate the propagation behavior of the model in a signalized intersection, similar to an urban road network, by comparing the expected theoretical results vs actual simulated results. The scope is restricted for this research towards signalized intersection, owing to its larger relevance in an urban setting.

- *Route choice (general):* The performance of the route choice submodule is evaluated by comparing the expected theoretical results vs actual simulated results.

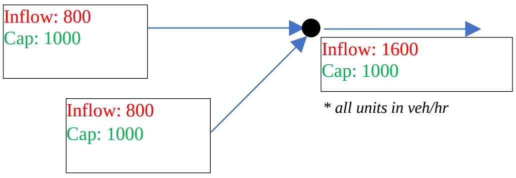
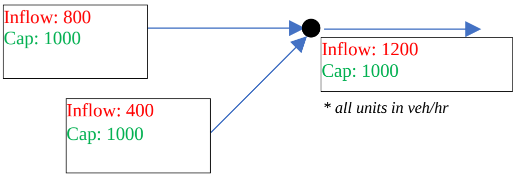
- *Route choice (route overlap)*: The performance of the route choice submodule along with the effect of route overlap is evaluated by comparing the expected theoretical results vs actual simulated results.

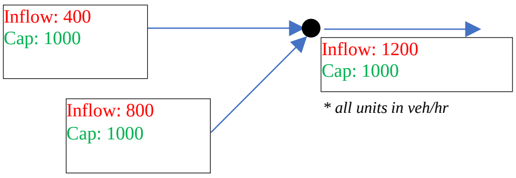
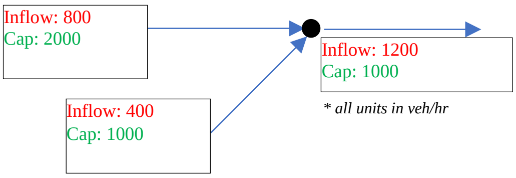
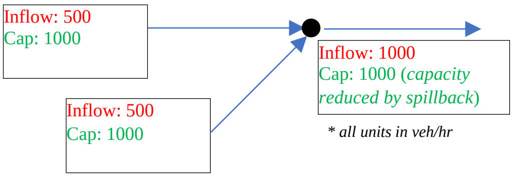
A series of 31 tests in 4 different categories have been identified for scoring these MoPs. The categorization of tests is based on the sub-components of a typical Macroscopic DTA model. The literature motivation and the theoretical test networks were mainly adopted from (FakhraeiRoudsari, Huang, & Tampère, 2015). The test series with descriptions are as provided in Table 3.

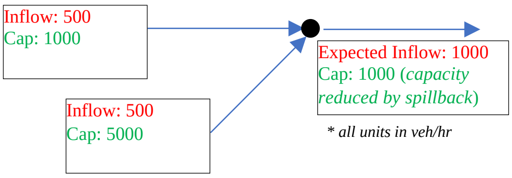
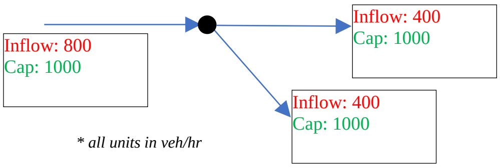
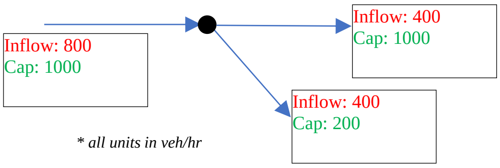
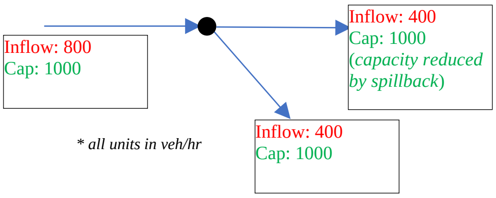
Table 3: Test series used to evaluate Tractability-based MoPs in EMMa

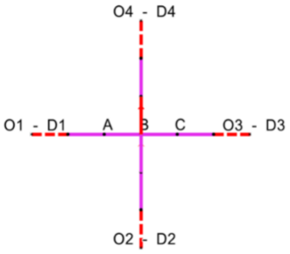
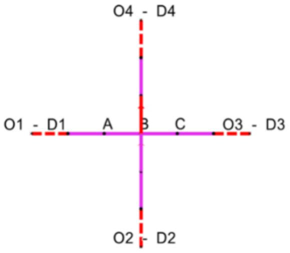
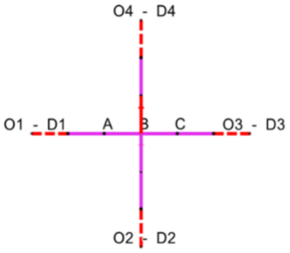
Sl No	Test Name	MoP Evaluated	Test Description and Expectations
1	**Flow propagation model - Output retrieval/ visualization**		**Test series to verify if the queue formation and dissipation are modelled correctly, i.e., to accurately identify the bottlenecks and to create the correct amount/severity of:** • **spillback upstream (spatial extent, speed/flow reduction)** • **flow reduction downstream** **For the entire test series, a single corridor network is used, with multiple links between one OD pair.**
1.1.1	Default speed parameters & connector speed @ 90Kmph	Propagation - Link flows	Simple demand propagation through a single network corridor, where link flows are in undersaturated condition throughout the simulation period.

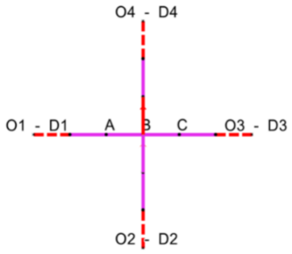
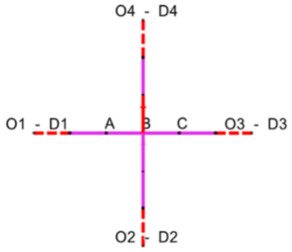
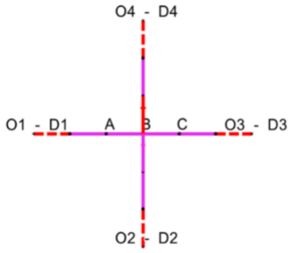
Sl No	Test Name	MoP Evaluated	Test Description and Expectations
1.1.2	Default speed parameters & connector speed @ 90Kmph, nue = 140	Propagation - Link flows	Influence of anticipation term in speed equation is checked through the test. A higher value of nue (parameter for anticipation term) would mean, the vehicles are more sensitive to link flows downstream so as to anticipate the speeds, when compared to other speed terms. The test is restricted to those DTA models with second order effects.
1.1.3	Default speed parameters & connector speed @ 50Kmph	Propagation - Link flows	Influence of connector speeds in the adjacent links is checked (connectors are links connecting the centroid and adjacent node). A lower connector speed would ideally mean that the propagation in the adjacent links will be slower, as a result of convection. The test is restricted to those DTA models with second order effects.
1.2.1	Free flow propagation with queue at the origin- connector speed @ 90Kmph	Propagation - Queuing	Demand propagation through a single network corridor. The simulation of demand is expected to create a bottleneck at the origin as the link is oversaturated (demand > capacity), in a specific time period of the simulation.
1.2.2	Fixed Bottleneck- Default speed parameters & connector speed @ 50Kmph	Propagation - Queuing	Influence of connector speeds in the adjacent links is checked, when the links are oversaturated. A lower connector speed would mean the that propagation in the adjacent links will be slower. The test is restricted to those DTA models with second order effects.
1.3.1	Fixed bottleneck with constant demand	Propagation - Queuing	A bottleneck is expected to form from the beginning of the simulation. Flow and speed characteristics of the links upstream and downstream of the bottleneck are analyzed (through propagation charts, space-time diagrams etc.) to understand the extent of queuing, spillback, capacity drop etc. Flow and speed distribution of a link upstream of the bottleneck is used to score quantitatively the curve roughness parameters as explained in Section 3.2.1.
1.3.2	Bottleneck with constant demand zero nue	Propagation - Queuing	Influence of anticipation term in speed equation is checked in Test No. 1.3.1, by reducing the value of the nue parameter to zero. The test is restricted to those DTA models with second order effects.

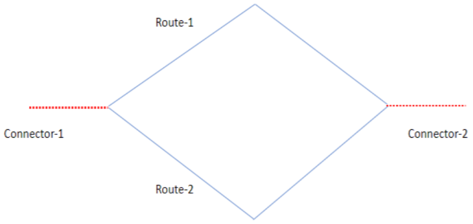
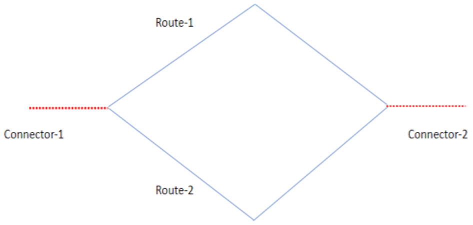
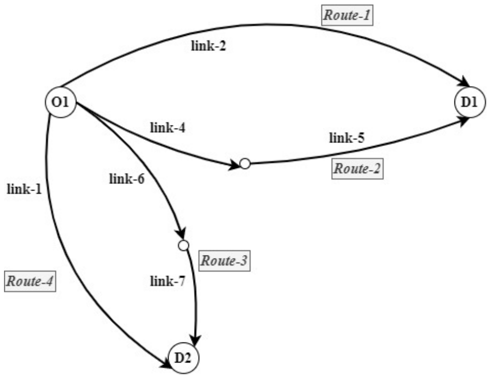
Sl No	Test Name	MoP Evaluated	Test Description and Expectations
1.4	Fixed bottleneck with peak demand	Propagation - Queuing	The propagation and queuing behavior of the links in case of peak demand is checked. Peak demand in this test refers to a specific period in simulation, where demand is greater than the capacity of the bottleneck link.
1.5	Bottleneck with constant demand variable capacity	Propagation - Effect of link-level traffic controls	The effect of dynamic link attribute editor is checked through this test. Essentially, a link attribute editor, alters the traffic states for a specific link for a specific period in time. This results in a variation in the fundamental diagram of the link. A control is provided in the bottleneck link which alters the saturation flow of the lanes in the link for a specific period. The variations in the propagation are further checked when the control is active.
1.6	Effect of variable speed limit in propagation	Propagation - Effect of link-level traffic controls	Similar to test no 1.5. Here the variation is for the free-flow speed of the link for a specific period in time. The control employed is expected to mimic variable speed limit in a real-world traffic network.
2	**Highway corridor with on-ramps and off-ramp (merges and diverges)**		**Test series to check the reliability in modeling congestion due to discontinuities in the network corridor. The role of merges and diverges are cross-checked (in addition to being potential primary bottlenecks): do they cause the right amount of spillback to the right upstream links, creating the right amount of delay?**
2.1.1	Single merge behavior receiving link's capacity is the constraint and **both inflow of sending links exceeds** their reduced outflow capacity	Node model-merge behavior	 The expectation here would be that the outflow of both the sending link reduces to half, so as to match the capacity of the receiving link. This reduction is bound to create queuing in the links upstream from the sending links. The northern links belong to the highway and the southern link is the on-ramp.
2.1.2	Single merge behavior receiving link's capacity is the constraint and **only highway inflow exceed** the reduced outflow capacity	Node model-merge behavior	

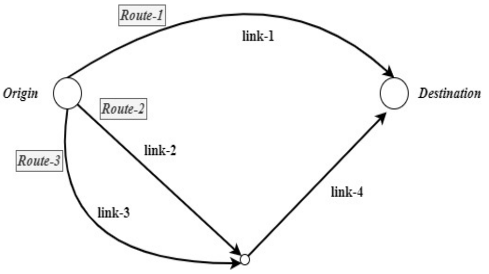
Sl No	Test Name	MoP Evaluated	Test Description and Expectations
			The expectation here would be that the outflow of both the sending link reduces to half, so as to match the capacity of the receiving link. This reduction is bound to create queuing in the links upstream from the <u>highway link</u>, where inflow exceeds reduced capacity.
2.1.3	Single merge behavior receiving link's capacity is the constraint and **only on-ramp inflow exceed** the reduced outflow capacity	Node model-merge behavior	 The expectation here would be that the outflow of both the sending link reduces to half, so as to match the capacity of the receiving link. This reduction is bound to create queuing in the links upstream from the <u>on-ramp link</u>, where inflow exceeds reduced capacity.
2.1.4	Single merge behavior receiving link's capacity is the constraint - **capacity of the highway merge link is twice that of the on-ramp link**	Node model-merge behavior	 The test checks the feature of capacity proportionality for merging links. For those merge model where blocking-back occurs on the basis of capacity of the sending links (StreamLine-MaDAM and StreamLine-eGLTM); the flow re-distribution upstream from the merge node is expected be proportional to the input capacity of the sending links. It may be noted that, for certain other merge models, where the re-distribution is proportional to other link characteristics such as no of lanes (MARPLE), this test would not yield different results.
2.2.1	Single merging behavior in the event of congestion, triggered by **spillback from a more downstream bottleneck.**	Node model-merge behavior	 Spillback from a bottleneck link downstream of the receiving merge link, is expected to reduce its capacity. This reduced capacity is further expected to reduce the outflow of both the sending

Sl No	Test Name	MoP Evaluated	Test Description and Expectations
			link, so as to match the capacity of the receiving link. This reduction is bound to create queuing in the links upstream from the sending links, where inflow exceeds reduced capacity.
2.2.2	Single merging behavior in the event of congestion, triggered by spillback from a more downstream bottleneck. - **capacity of the sending links are different**	Node model-merge behavior	 This test checks the combination of test no 2.1.4 and 2.1.1, with expectations similar to those tests.
2.3.1	Simple diverge model under free-flow conditions	Node model-diverge behavior	 Free-flow propagation is expected in all the links of this test network, as throughout the simulation, the links are in under-saturation condition.
2.3.2	Simple diverge model **when the capacity of the receiving link is the constraint**	Node model-diverge behavior	 The outflow of the diverge link in the highway (receiving link) is expected to reduce to the tune of the capacity of the diverge link in off-ramp, to ensure continuity. The northern links belong to the highway and the southern diverge link is off-ramp.
2.3.3	Simple diverge model **due to congestion and a spillback from a bottleneck downstream of the diverge node**	Node model-diverge behavior	 Expectation similar to test 2.3.1. The capacity reduction in highway diverge link (due to spillback from bottleneck downstream) is expected to reduce the outflow of the diverge link in off-ramp.

Sl No	Test Name	MoP Evaluated	Test Description and Expectations
2.4.1	Consistency of turning rates at the diverge in comparison with the O-D demand	Node model-diverge behavior	The test is performed to check if the restriction imposed at the nodes due to capacity constraints in the link affects the total vehicles reaching the destination from the origin.
3	**Signalized intersection**		**Delays in the urban network due to vehicular intersections. Influence of delay in route travel time.**
3.1.1	Intersection behavior in undersaturated conditions	Signalized Intersection	 Demand between O1 and D3 propagated for the simulation period, with all links in between in **undersaturated** conditions. Effect of delay caused in the network due to signal at B is checked. No delay is expected at the intersection, as the links are in undersaturated condition.
3.1.2	Intersection behavior in oversaturated conditions	Signalized Intersection	 Demand between O1 and D3 propagated for the simulation period, with all links in between in **oversaturated** conditions. Effect of delay caused in the network due to signal at B is checked. A delay is expected at the intersection, as the links are in oversaturated condition.
3.1.3	Modeling intersections - Single Turn Behavior-Spillback from down stream	Signalized Intersection	 Demand between O1 and D3 propagated for the simulation period. Link CD3 is the bottleneck as the demand is greater than capacity. Effect of delay caused in the network due to signal at B is checked. A delay is expected at the intersection, as the bottleneck link in CD3, causes spillback upstream and thereby results in capacity reduction.

Sl No	Test Name	MoP Evaluated	Test Description and Expectations
3.2.1	Modeling the intersections - Diverge Behavior	Signalized Intersection	 Demand between O1D2, O1D3 and O1D4 is propagated to check the diverge behavior at intersection B. All links are in **undersaturated** condition throughout the simulation. No delay is expected at the intersection, as the links are in undersaturated condition.
3.2.2	Modeling intersections - Diverge Behavior in Oversaturated conditions	Signalized Intersection	 Demand between O1D2, O1D3 and O1D4 is propagated to check the diverge behavior at intersection B. All links between O1 and D3 are in **oversaturated** condition throughout the simulation. A delay is expected at the intersection, as the links are in oversaturated condition.
3.2.3	Modeling intersections - Diverge Behavior during spill back from downstream link	Signalized Intersection	 Demand between O1D2, O1D3 and O1D4 is propagated to check the diverge behavior at intersection B. Link CD3 is the bottleneck as the demand is greater than capacity, which is expected to cause spillback. A delay is expected at the intersection, as the bottleneck link in CD3, causes spillback upstream and thereby results in capacity reduction.
4	**Route Choice**		**Importance and relevance of Stochastic Dynamic User Equilibrium**
4.1.1	Modeling the Route Choice - Simple Route choice with demand < capacity	Route choice (general)	A two-route network between a pair of origin and destination is used for understanding the route choice behavior. A stochastic MNL assignment is tested with MSA scheme for averaging the route costs between the iterations. All links are in undersaturated conditions. Expectation would be equal route choice proportions in both the routes, for all the route choice moments and/or iterations.

Sl No	Test Name	MoP Evaluated	Test Description and Expectations
4.1.2	Modeling the Route Choice - Simple Route choice when cost of one route slightly more than other (Undersaturated)	Route choice (general)	 The route choice behavior is evaluated when the travel cost of route-1 is lesser than route 2. A larger route choice proportion will be expected for the cheaper route, as all the links are in undersaturated conditions, throughout the simulation.
4.1.3	Modeling the Route Choice - Simple Route choice when cost of one route slightly more than other (Oversaturated)	Route choice (general)	 Route costs are similar to previous test. During the second hour of the simulation, the links in Route-1 experience oversaturation, and the route choice variations are evaluated. The route choice proportions for the cheaper route are expected to reduce, when the links get oversaturated. This is due to the expected increase in route cost as a consequence of congestion and delay in the oversaturated route.
4.2.1	Route choice - Independent routes between 1 origin and 2 destinations	Route choice (general)	 Route choice behavior for a larger network is studied. Travel time costs of links prior to simulation start are the same. The expectation of route choice and route proportions for all four routes should be in-line with the saturation conditions of the links, as the initial travel times of all the links are maintained equal.

Sl No	Test Name	MoP Evaluated	Test Description and Expectations
4.2.2	Relative difference in route proportions - MNL vs PCL/C-Logit, and consequent influence of route overlap	Route choice (route overlap)	 The network is used to study the influence of route overlap in route choice. Link-4 is the overlap link in the southern routes – route 2 and 3. The travel time on all routes are same with and all links are in undersaturated conditions throughout the simulation period. A comparative study is conducted between MNL proportioning and PCL/C-Logit proportioning to understand this difference. The expectation would be that route choice iterations which accounts for route overlap such as PCL, C-Logit, should account for the independence of route 1 and distribute more traffic to the independent route. The expectation is motivated from (Chen, Kasikitwiwat, & Ji, 2003). The simulations are run as One-shot, as the effect of averaging scheme and iterations is not of relevance in this test.

3.2.5. Integration of Network Hierarchies - Urban and Motorway Roads

The ability for a Macroscopic DTA model to mimic the driving condition of an urban road and a motorway road is evaluated through this MoP. The evaluation is conducted qualitatively as per the scoring system provided in Table 2. The fluctuations of traffic states (mainly traffic speeds) over a series of link are studied to evaluate the behavior.

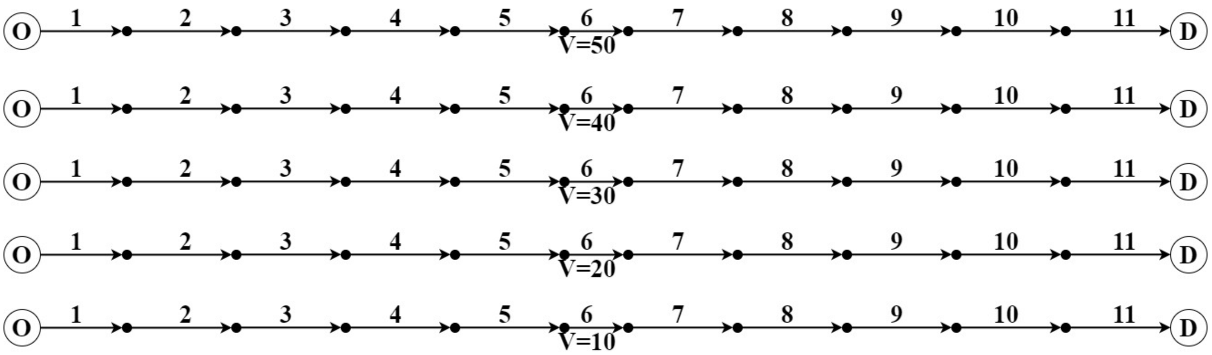


Figure 9: Test network used for evaluating the speed fluctuations in urban and non-urban links

The test involves a comparative study between a series of urban links and non-urban links (highway links), based on the assumption that in urban roads, driver behavior is more aggressive and speed fluctuations are more sudden (Raadsen, Mein, Schilpzand, & Brandt, 2010). The test network involves the propagation in single network corridor between a pair of origin-destination with constant demand. The network used for the test is as shown in Figure 9. The free-flow speeds in all links except Link No.6 is fixed at 50 Kmph. Over a series of 5 networks, the free-flow speed for Link No.6 is reduced from 50 Kmph to 10 Kmph. Length of all links except Link No. 6 is 300m. Length of Link No. 6 is 20m. Demand between O and D is simulated in undersaturated conditions for all the links.

Two separate test scenarios are analyzed for urban link specification and non-urban link specification. As the model characteristic under evaluation is dynamic speed fluctuation, the test is mainly relevant for a second-order model that can incorporate hysteresis. The MoP name for this category is *Fluctuation of traffic states over a series of urban and non-urban links.*

3.2.6. Computational Efficiency

Feasibility of model run is an important characteristic of a macroscopic DTA model that needs to be evaluated for any application type of model user type. The 2 MoPs are used in EMMa for evaluating the same, which are measured quantitatively as listed below:

- *Run Time in Sec*: The MoP measures the time taken for completing the simulation for a specific test network run with the Macroscopic DTA model under evaluation. It is important to run the simulations in the same network for all the DTA models under evaluation, run with the same computer hardware. It may be noted none of DTA's used in this research has a multicore implementation. This motivates the direct usage of run time in seconds.

- Peak Memory Usage: The MoP measures the peak computer memory used during the simulations of the model. The peak memory is usually witnessed in the final iteration before convergence. Similar to the previous MoP, the same test network is used for all the DTA models under evaluation, run in the same computer hardware.

3.2.7. Usability

Another important feature for evaluating any macroscopic DTA model is the ability to use the model seamlessly. Departing with the definition and literature context provided in (Seffah, Donyaee, Kline, & Padda, 2006), 7 MoPs are used in EMMa for evaluating usability as listed below:

- *Familiarity*: Whether the user interface offers recognizable elements and interactions that can be understood by the model user.

- *Simplicity*: Whether extraneous elements are eliminated from the user interface without significant information loss.

- *Navigability*: Whether model users can move around in the model windows in an efficient way.

- *Controllability*: Whether model users feel that they are in control of the DTA model.

- *Readability*: Ease with which visual content (e.g., text dialogs) can be understood.

- *User guidance*: Whether the user interface provides context-sensitive help and meaningful feedback when errors occur.

- *Flexibility*: Whether the user interface of the DTA model can be tailored to suit model users' personal preferences.

All of the above MoPs are measured qualitatively by means of the score system provided in Table 4.

Table 4: Qualitative scoring system to measure Usability

Score from	Score to	Description	Usability Measure
1	10	1- The model interface is complex for a regular user to understand, and would require external help. 10 - The model interface is straightforward and can easily be operated.	Familiarity
1	10	1- The model is too sophisticated with a lot of irrelevant features, which does not affect the expected results of the model. 10 - The model working is effortless and contains hardly any non-relevant parameters or features.	Simplicity
1	10	1- The model interface is complex and movement from one window to the other is demanding 10 - Navigation between the interface windows and tools is fluid and user-friendly	Navigability
1	10	1- The modeling environment is strict and does not offer the user the required amount of transparency and control 10 - The model is completely transparent and the model user feels completely in control of the working of the model.	Controllability
1	10	1- Visual readability is poor and the model user requires prior practice and external help to use the tools 10 - The dialogue boxes, plug-ins and the tools in the model are straightforward, clear and consistent	Readability
1	10	1- The buttons and features in the interface are not descriptive enough for the model user to explore. The error log can only be understood by a model developer or an experienced programmer. 10 - The user interface provides the appropriate help texts and prompts to use the tools in the model. The error log is descriptive enough for a new user to engage in debugging.	User guidance
1	10	1- The User interface is inelastic and does not allow any form of alterations as per the preferences of the user 10 - The User interface can be easily customized according to the preferences of the model user	Flexibility

It is interesting to note that for scoring "Usability" qualitatively, the range provided is more disaggregate compared to the score system defined in Table 2. The larger gap in the range is bound to incorporate the variations in scores for a commercial model developed with a larger focus on interface aesthetics and user experience.

3.3. Secondary Dimensions in EMMa

MoPs form the core of evaluation in EMMa, by providing the modeler the ability to score the models. The MoPs thus create a tabular score matrix for the DTA model under evaluation. In EMMa, other than the score value, each MoP is also governed by weights which differs on the basis of some secondary dimensions. These other dimensions, which bring alterations to the outcome of the scores of the MOPs are as listed below:

- **The Model User Type**
    - *Policy Maker*: A policy maker in EMMa refers to a transport professional working in the public sector (governmental agencies), who is directly or indirectly involved in the transport-related decision-making processes in the government through the usage of traffic models.
    - *Mobility Consultant*: A mobility consultant refers to a transport professional working in an engineering consultancy, who is involved in providing solutions to transport-related problems by applying and analyzing them virtually in simulation-based traffic models.
    - *Scientific Researcher*: A scientific researcher in EMMa, refers to an expert in traffic modeling and simulation, who has relevant research experience specifically in DTA model application. He/she uses simulation-based traffic models for various research themes.
    - *Model Developer*: A model developer would primarily be a scientific researcher in the field of simulation-based traffic models who has attained experience not only in applying and using DTA models, but also in developing them as a software product.
- **The Application Planning Horizon**
    - *Strategic Planning*: Refers to a decision-making context where analysis and choices have major systemwide and long-term impacts, and usually involve resource acquisition and network design (Ortúzar & Willumsen, 2011). The decision can have a time horizon of **5 or more years**. The solutions involved can include construction of additional infrastructure-roads, bridges, transport lines, bike lanes etc., which requires planning, execution and operation over a long period of time.
    - *Tactical Planning*: Decisions based on tactical planning attempts to solve issues having a narrower perspective and concern questions like making the best use of existing facilities and infrastructure (Ortúzar & Willumsen, 2011). The decision can have a time horizon of **1 to 5 years.** The solutions involved can include traffic management, traffic re-routing plans, introducing toll in lanes (including operations of High-Occupancy Toll (HOT) lanes and High-Occupancy Vehicle (HOV) lanes), increasing the frequency of bus trips in a route by diverting buses etc.
    - *Operational Planning*: Decisions for operational planning, involve the narrowest time horizon which can vary from **weeks/days to even real-time**. Some examples where operational planning is involved are: road works in urban and non-urban roads, real-time

re-routing of buses based on optimized routes, disruption in roads due to accidents – Incident management etc.

3.4. Determining the Weights for Evaluation

Secondary dimensions are incorporated in EMMa by influencing the weights of the MoP scores. As part of developing EMMa, these weights were obtained by conducting surveys with experts in each of the model user category. The following steps were carried out by the researcher for conducting the Model User survey for EMMa:

1. Design the questionnaire for the experts to provide weights.

2. Identify the expert panel of model user.

3. Conduct interviews with the panel to understand the motivation behind the responses.

4. Calculate the average weightage per model user type and per application horizon for each of the MoPs.

Designing the questionnaire, involved the task of linking the responses of the experts to the MoPs in EMMa. Furthermore, the design of the questions should also factor the knowledge and expertise of the respondent in traffic theory and traffic modeling. In this regard, two sets of questions were designed for 1) Policy Maker and Mobility Consultant 2) Scientific Researcher and Model developer. The question set and their corresponding link to the MoPs are as illustrated in Table 56 and Table 57 in Appendix-B. Each of the questions were provided with a response matrix with the option to score for each application horizon, as shown in Figure 173, Appendix-B.

Owing to the constraint of time and availability of experts and the researcher, a total of 10 model users across various user categories, primarily based in Netherlands were identified. The expert panel included – 3 mobility consultants, 2 Public sector professionals and 5 Research professionals. As some of the experts belonged to multiple model user categories, the total number of expected responses were 15. The break-up of the expected responses was as follows: *3 X Policy Maker*, *3 X Mobility Consultant, 5 X Scientific Researcher*, *4 X Model Developer*.

Of the 10, 7 users responded to the survey request and a total of 12 responses were obtained (*2 X Policy Maker*, *3 X Mobility Consultant*, *3 X Scientific Researcher*, *4 X Model Developer*). Initially, the number of years of experience of the experts in the panel were noted, to use the same for weighted averaging of the responses. However, as all the experts had 10+ years of experience, the weights responded by the panel were averaged out arithmetically, which ultimately became an input for EMMa.

Summary of Weights

The survey responses are linked to each MoP and their categories as per Table 56 and Table 57, Appendix-B. The weights given by each model user is arithmetically averaged per category, which serves as an input for EMMa. The responses have further been categorized based on the application horizon.

It can be noted that for **Strategic Planning** (Figure 10**Error! Reference source not found.**), in general the *Model Robustness* has been provided with a high weightage, which is followed by *Applicability* and *Tractability* among all the model user categories. As robustness describes the ability of the model to handle traffic assignment and provide a stable result, for strategic planning application, it is weighed with a high priority by all the model users, which is as expected. Equilibrium is of utmost importance in a strategic planning application, as the model results are looked at network level. Upon discussion with Scientific researchers, it was understood that there is a growing research interest for improving the robustness of the models, which puts the MoP under spotlight. As long-term planning is involved, the ability of the model to incorporate multiple classes also becomes very important. For a Policy Maker, the *Computational Efficiency* and the *Usability* scores are low which can be expected as he/she may not be directly involved in the application of the model, but is rather interested in the results and interpretation of the model results.

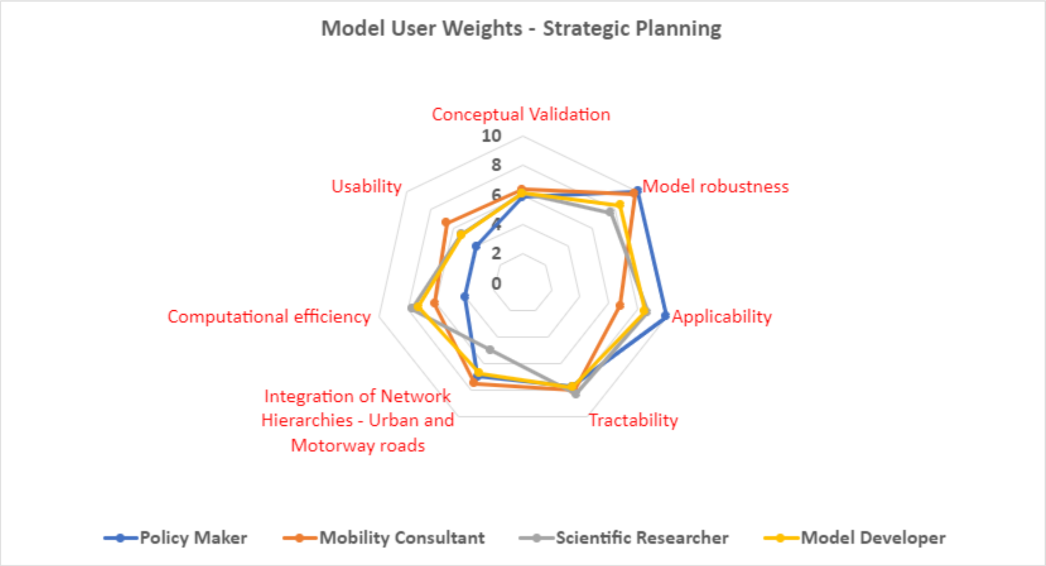


Figure 10: Model User weights for each MoP Category - Strategic Planning Application

Computational Efficiency is provided a high value by the model users directly involved in the development and the research of DTAs – Scientific researcher and Model Developer. Interestingly the model *Usability* becomes a priority for the Mobility Consultant, as he/she may be involved in frequent use of the model for providing solutions to transport-related problems involving large-scale networks, and may not be familiar with core traffic modeling. This forms their need for the DTA model to be easily used and applied to various traffic networks within feasible model runs.

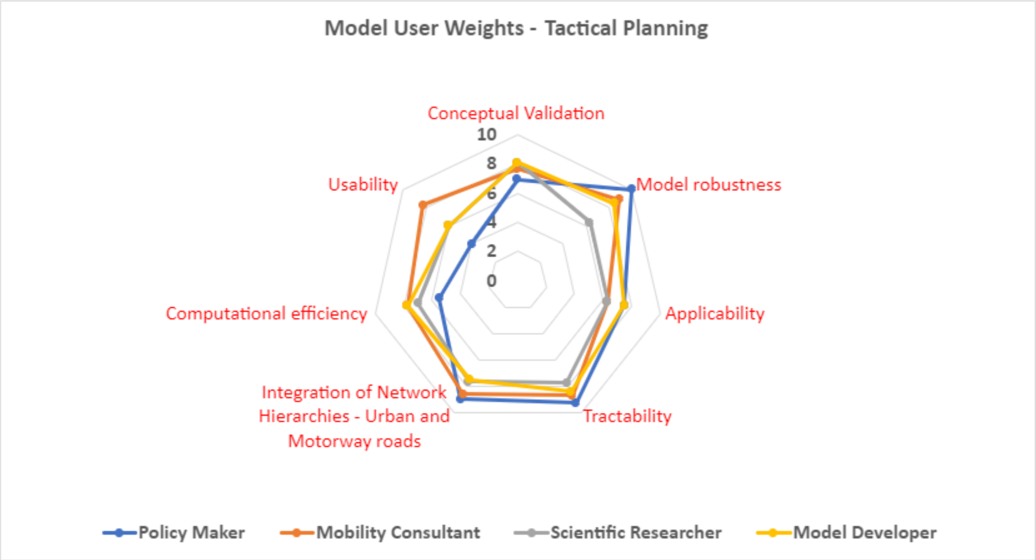


Figure 11: Model User weights for each MoP Category - Tactical Planning Application

Results for **Tactical Planning** (Figure 11) are similar to that of strategic planning, with a high weightage for *Model Robustness* among all model user categories, except for scientific researcher. We observe a growing importance of the theoretical *Tractability* of the model results when the time horizons get more disaggregated. The importance of *Applicability* in tactical planning has decreased among all model users especially for Policy Maker, compared to strategic planning. This can be expected as the intended use of tactical models are for a specific transport solution such as rerouting plans or use of peak-hour lanes etc., which are in most cases mode-specific and user-class specific.

Similar to the trend observed in *Tractability*, we do observe an increase of importance for *Conceptual Validation*, especially among Scientific Researcher and Model Developer. The need for the model to handle phenomenon observed in real-world situations are emphasized by the model users especially when the application time horizon gets more disaggregated for Tactical planning. The emphasis on model *Usability* is stressed by the Mobility Consultant even for tactical planning.

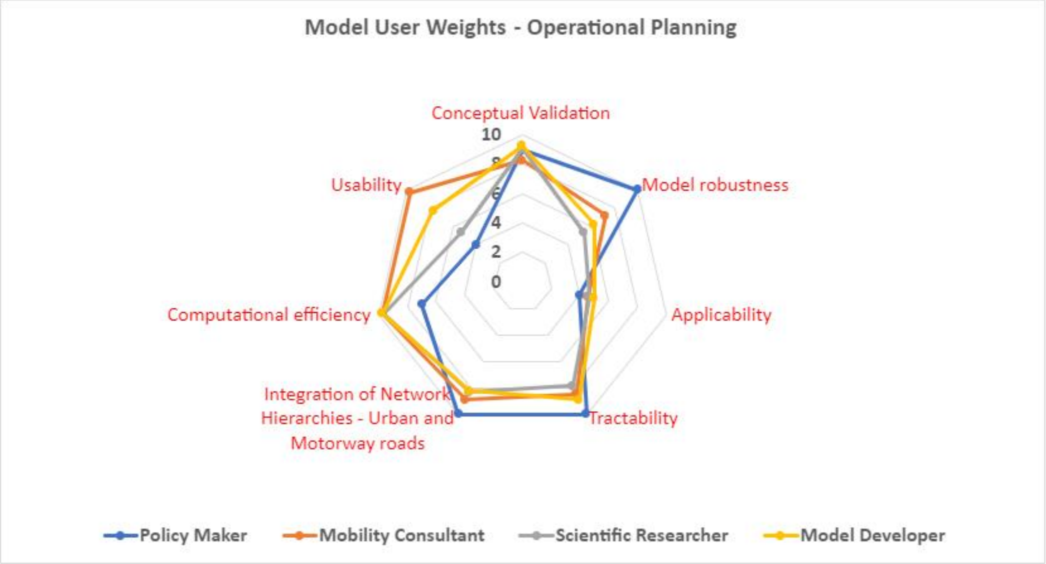


Figure 12: Model User weights for each MoP Category - Operational Planning Application

From the results of **Operational planning** (Figure 12), it can be understood that the importance of *Model robustness* has decreased substantially amongst all model users. Upon discussion with the experts, it was understood that the intended use of the operational model was for a real-time application and at a finer spatial level of analysis, with a smaller network (part of an urban road or neighborhood). In such a case, route choice and *Model Robustness* does not play a major role when compared to demand loading and propagation behavior. This is because in most cases, the route choice and route fractions are calculated pre-trip and the focus of analysis will be shifted to the behavior at a link level or a node level. This shift of preference is effectively captured by a lower score of *Model Robustness* by Mobility consultant, Scientific researcher and Model developer. A variation is observed for a Policy Maker perspective, who still believes in the importance of an accurate route choice for operational planning purposes.

The theoretical *Tractability*, which can be roughly translated to the effect of the propagation model, is scored very high by all model users. The emphasis here again would be the sensitivity of the model to behavioral changes at the link level. This may be to study the effect of say, incidents or traffic management controls and the subsequent changes in the traffic flow. At an operational level, it becomes imperative that the model is capable in incorporating these behavioral effects such as capacity drop, spillback effects, node-model accuracy etc. The quickness of the model run in terms of *Computational Efficiency* is given high importance for by all model users, as for a real-time operational planning application, it is essential that the results are obtained as quick as possible to make real-time decisions. For the MoP category – *Integration of network hierarchies – urban and Motorway road*, the weightage allotted by a Policy Maker increases substantially from strategic→operational planning. Upon discussion with experts, it was understood that they do not expect a seamless modeling of urban and motorway network propagation in a strategic planning case, looking at the network level. However, for an operational planning application, this speed variation based on network hierarchy plays a huge role in the behavior of traffic at the road, because of which a high weightage is given.

It is also interesting to note that a Model Developer puts a high weight on *Usability* in operational planning application. A model developer has an added motivation to increase the usability of the model especially for operational planning applications, to reach a larger group of model users, such as a municipality, or traffic authority, thereby enriching the model application and validation even at an entrepreneurial perspective.

3.5. EMMa – Model Working

This section will cover in detail the working of EMMa, describing the interplay of the dimensions and various techniques which will provide the modeler the ranking of the models that are being evaluated. As stated before, the rankings vary per application type and model user type, which are provided as choice inputs by the modeler. The flow chart describing the working is as shown in Figure 13.

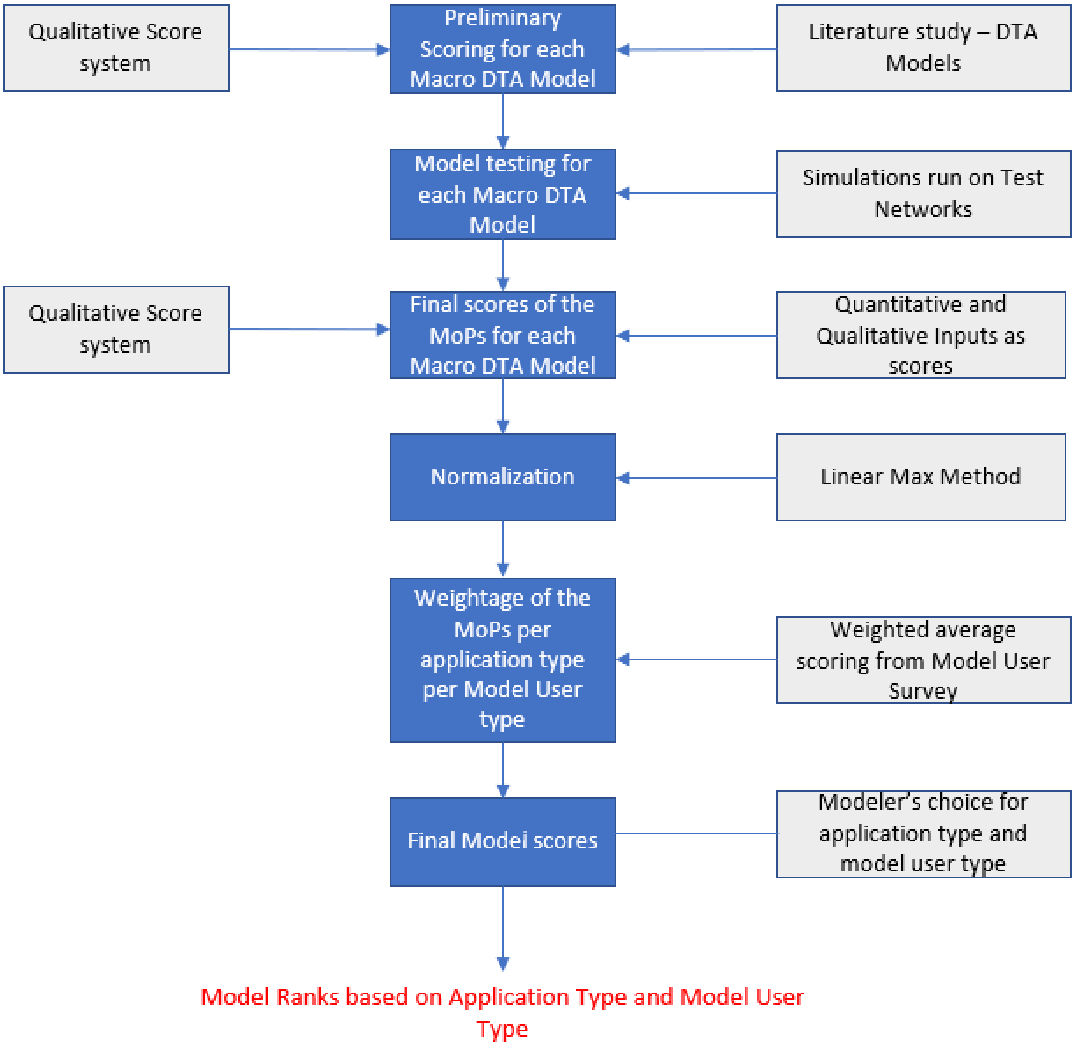


Figure 13: Step-by-step working of EMMa with inputs

As mentioned before, the first step in applying EMMa is to study the theoretical working of the DTA models under evaluation. Literature study on the same will help the modeler to devise the preliminary scores. The next step would to be run the test networks for each model under evaluation, as per Table

3. It may be noted that the MoPs listed in EMMa are extensive. The framework provides the option for the modeler to use those MoPs relevant for his/her evaluation, if needed. After running the tests, the modeler will be in a position to fill the scores of the DTA models under each of the MoPs. The next step would be to standardize all the different MoP scores, to a consistent score matrix. This step is important to provide an apple-to-apple comparison between the MoPs, as certain quantitative scores are better off, when the score value in itself is low (eg. simulation run time, peak memory usage etc.).

A number of normalization techniques were available from literature (Binsbergen, 2020), which were tested out individually before finalizing for EMMa, results of which have been summarized in Figure 14. The techniques tested are as listed in Table 5.

Table 5: Normalization techniques tested in EMMa, adopted from (Binsbergen, 2020)

Sl No	Normalization Technique	When higher score is better	When lower score is better
1	linear: max (all related to 1)	$n_{ij} = \frac{r_{ij}}{r_{max}}$	$n_{ij} = 1 - \frac{r_{ij}}{r_{max}}$
2	linear: max-min, non-proportional (spread 0-1; std interval)	$n_{ij} = \frac{r_{ij} - r_{min}}{r_{max} - r_{min}}$	$n_{ij} = \frac{r_{max} - r_{ij}}{r_{max} - r_{min}}$
3	linear: sum (sum = 1)	$n_{ij} = \frac{r_{ij}}{\sum_{j=1}^{m} r_{ij}}$	$n_{ij} = \frac{1/r_{ij}}{\sum_{j=1}^{m} 1/r_{ij}}$
4	Vector ('Euclidian')	$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{j=1}^{m} {r_{ij}}^2}}$	$n_{ij} = 1 - \frac{r_{ij}}{\sqrt{\sum_{j=1}^{m} {r_{ij}}^2}}$
5	Logarithmic (sum = 1)	$n_{ij} = \frac{\ln(r_{ij})}{\ln(\prod_{j=1}^{m} r_{ij})}$	$n_{ij} = \frac{1 - \frac{\ln(r_{ij})}{\ln(\prod_{j=1}^{m} r_{ij})}}{m - 1}$

$$r_{ij} = non\ normalized\ value\ of\ MoP\ score\ \text{"i"}\ for\ model\ \text{"j"}$$
$$n_{ij} = normalized\ value\ of\ MoP\ score\ \text{"i"}\ for\ model\ \text{"j"}$$

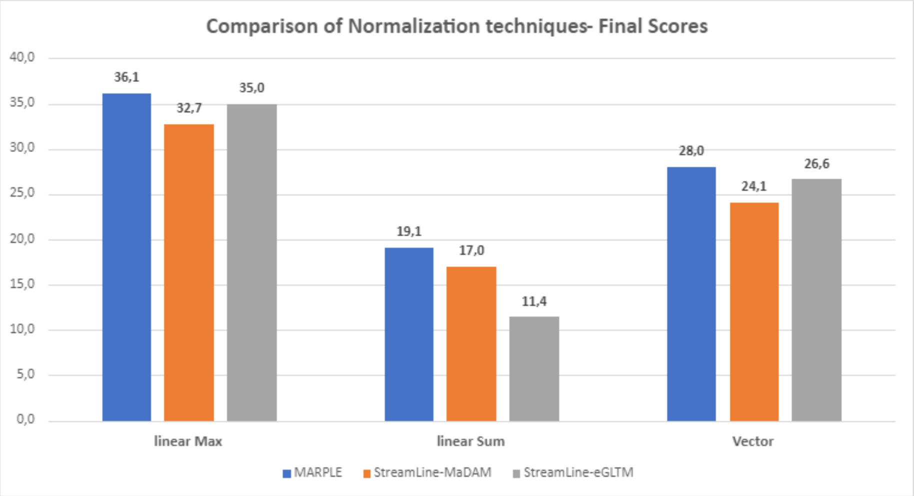


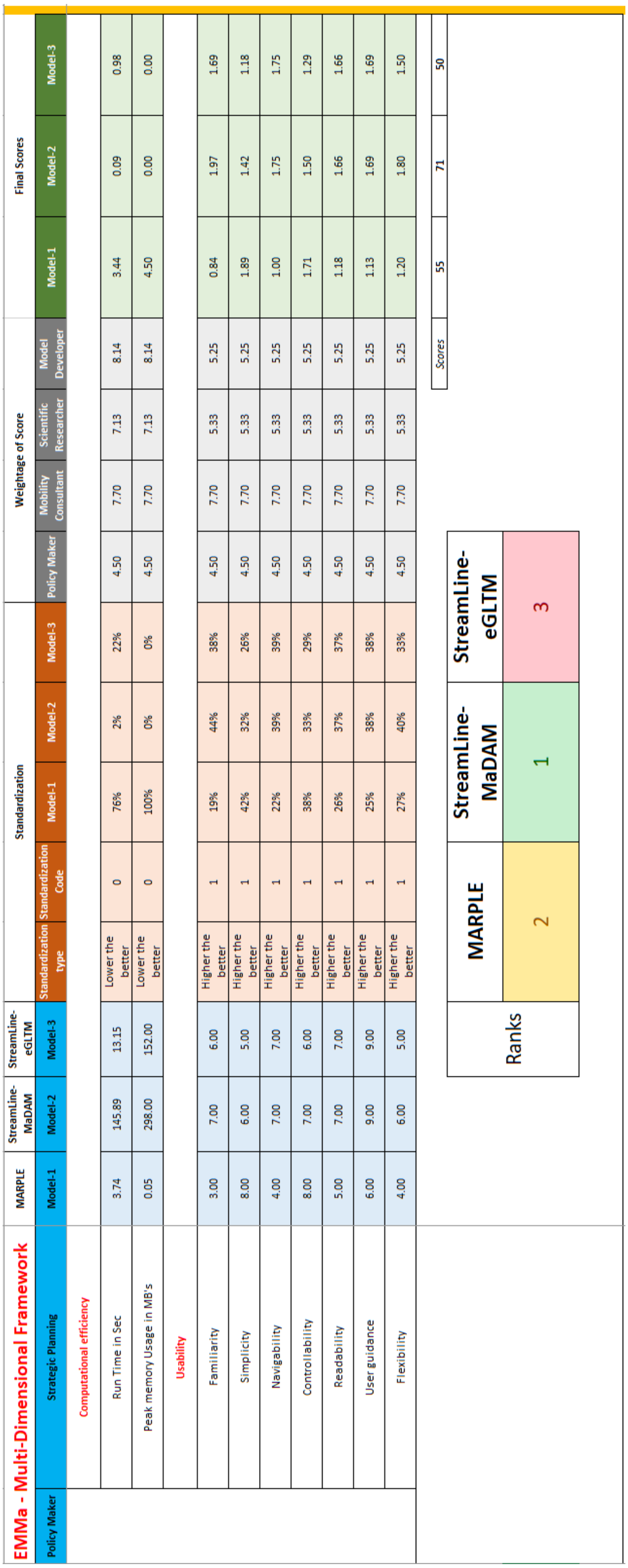
Figure 14: Comparison of results – Normalization techniques

From all the normalization techniques, the Linear: Max method is adopted in EMMa. As suggested from the literature, the selection of the technique involved testing of the results from all the methods stated in Table 5, as shown in Figure 14. Both Linear: Max-Min and Logarithmic are eliminated from sensitivity test as they standardized the evaluation score to a zero value for all three models when their initial scores were same (which might be the case for binary measurement), which is wrong. As observed in Figure 14, the Linear: Sum is highly sensitive to small change in scores. For example, in case of model robustness MoP, even though MARPLE and StreamLine: eGLTM had minor differences in initial score, eGLTM obtained a substantially smaller final score which is not right. When compared to Vector method, the Linear: Max method proved to be the most sensitive, when it comes to the final scoring. The method was also able to deal with zero scores without causing error values. This further motivated the choice of Linear: Max method for standardization.

After the standardization, the final scores per model per MoP, is obtained by multiplying the weights to the standardized scores and summing the average score for each of the 7 MoP categories. As stated before, the weights vary based on the input provided by the modeler for the application type and model user type. On the basis of the final score of the models, the ranking takes place and the modeler obtains the best macroscopic DTA model, suitable for a given application horizon and a given model user type.

The user interface of EMMa is as shown in Figure 15. The current version is developed in MS Excel, owing to it advantage of easy integration and accessibility. The “cells” indicated in blue, is the input provided by the modeler, based on which the final scores vary, thereby altering the final ranks.

The following section would describe the results from applying EMMa for the three DTA models, which was discussed in detail in the preceding section. The application will also become a validation of the evaluation framework to understand its usefulness and potential. Results from the theoretical test cases described in Table 3, will also be discussed in this section for each of the three models.



EMMa - Multi-Dimensional Framework		MARPLE	StreamLine-MaDAM	StreamLine-eGLTM	Standardization					Weightage of Score				Final Scores		
Policy Maker	Strategic Planning	Model-1	Model-2	Model-3	Standardization type	Standardization Code	Model-1	Model-2	Model-3	Policy Maker	Mobility Consultant	Scientific Researcher	Model Developer	Model-1	Model-2	Model-3
	Computational efficiency															
	Run Time in Sec	3.74	145.89	13.15	Lower the better	0	76%	2%	22%	4.50	7.70	7.13	8.14	3.44	0.09	0.98
	Peak memory Usage in MB's	0.05	298.00	152.00	Lower the better	0	100%	0%	0%	4.50	7.70	7.13	8.14	4.50	0.00	0.00
	Usability															
	Familiarity	3.00	7.00	6.00	Higher the better	1	19%	44%	38%	4.50	7.70	5.33	5.25	0.84	1.97	1.69
	Simplicity	8.00	6.00	5.00	Higher the better	1	42%	32%	26%	4.50	7.70	5.33	5.25	1.89	1.42	1.18
	Navigability	4.00	7.00	7.00	Higher the better	1	22%	39%	39%	4.50	7.70	5.33	5.25	1.00	1.75	1.75
	Controllability	8.00	7.00	6.00	Higher the better	1	38%	33%	29%	4.50	7.70	5.33	5.25	1.71	1.50	1.29
	Readability	5.00	7.00	7.00	Higher the better	1	26%	37%	37%	4.50	7.70	5.33	5.25	1.18	1.66	1.66
	User guidance	6.00	9.00	9.00	Higher the better	1	25%	38%	38%	4.50	7.70	5.33	5.25	1.13	1.69	1.69
	Flexibility	4.00	6.00	5.00	Higher the better	1	27%	40%	33%	4.50	7.70	5.33	5.25	1.20	1.80	1.50
													Scores	55	71	50

	MARPLE	StreamLine-MaDAM	StreamLine-eGLTM
Ranks	2	1	3

Figure 15: EMMa - User Interface

4. Results and Discussion

The application of EMMa to the three models under study will be illustrated and discussed in detail in this section. The chapter will first include the details on MoP-wise scoring for each of the DTA model under evaluation (Section 4.1). The second part of the chapter will show the results from scalability test, which is an essential step in evaluating some MoP's (Section 4.2). The next part of results involves the sensitivity analysis in the ranking of the models by varying the weights on the basis of application horizon and model user categories (Section 4.3). Interested readers are also referred to the working model of EMMa, attached as an appendix which can be read in conjunction with the report to experience the working of the tool as explained in Section 3.5. An important aspect while performing the simulation tests are the versions of the DTA models used for evaluation and the hardware used for testing. The reader may please note that the evaluation scores for the MoPs are based on these versions of the models and any improvements made post these versions are not be accounted while evaluating and scoring.

Hardware used for simulation testing: *Intel(R) Core(TM) i5-8265U CPU @ 1.60 GHz 1.80 GHz, with 8.00 GB RAM.*

StreamLine: MaDAM and StreamLine: eGLTM

OmniTRANS Version: 8.0.30.11960
StreamLine Version: 8.0.30

MARPLE

Version: 3.5.3 of 14-06-2021
OmniTRANS with MARPLE plugin Version: 8.1.505.12144

4.1. MoP Scores in EMMa

As suggested in the methodology, the first step in applying EMMa is to perform the literature review on the models and the underlying theoretical working of each of its sub-modules. The results of the same is shown in Section 2.3.3. This step is essential for the scoring some of the MoPs, which are binary measurements. Moreover, it provides the modeler the theoretical construct of the DTA model. The rest of the MoPs are scored on the basis of conducting tests as mentioned in Table 3. Appendix-AThe summary of the MoPs scores is as provided in Table 6.

Table 6: Summary of MoP scores in EMMa

Measures of Performance (MOP)	MARPLE	StreamLine-MaDAM	StreamLine-eGLTM
Conceptual Validation			
Flow Metering or Blocking back - strict capacity constraint	2.00	3.00	3.00
Traffic Spillback - strict storage constraint	3.00	4.00	3.00
Capacity drop	0.00	2.00	0.00
Link-level dynamic distribution of vehicle speeds- Curve roughness factor	44.57	33.64	34.78

Measures of Performance (MOP)	**MARPLE**	**StreamLine-MaDAM**	**StreamLine-eGLTM**
Link-level dynamic distribution of traffic flows- Curve roughness factor	18.39	1087.74	1089.20
Presence of variable route set	0.00	0.00	0.00
Modeling of stop and go waves	0.00	0.00	0.00
Model robustness			
Relative change in Gap Function/ Flows between final iteration	0.0002	0.0786	0.0009
Applicability			
Difference in Network Supply based on Modes	1.00	1.00	1.00
Difference in Input parameters based on different trip purposes	1.00	1.00	1.00
Tractability			
Propagation - Link flows	3.00	3.00	4.00
Propagation - Queuing	2.33	3.40	3.33
Propagation - Effect of time variability in Fundamental Diagram	3.50	2.50	0.00
Node model-merge behavior	2.83	2.83	4.00
Node model-diverge behavior	2.50	4.00	4.00
Signalized Intersection	1.83	2.67	2.17
Route choice (general)	4.00	3.00	3.00
Route choice (route overlap)	4.00	4.00	4.00
Integration of Network Hierarchies - Urban and Motorway roads			
Fluctuation of traffic states over a series of urban and non-urban links	0.00	4.00	0.00
Computational efficiency			
Run Time in Sec	7421.40	668989.09	9488.28
Peak memory Usage in MB's	712.78	1768.00	3992.00
Usability			
Familiarity	3.00	7.00	6.00

Measures of Performance (MOP)	MARPLE	StreamLine-MaDAM	StreamLine-eGLTM
Simplicity	8.00	6.00	5.00
Navigability	4.00	7.00	7.00
Controllability	8.00	7.00	6.00
Readability	5.00	7.00	7.00
User guidance	6.00	9.00	9.00
Flexibility	4.00	6.00	5.00

Please note that majority of the MoPs are based on a series of tests as mentioned before. The motivation of scoring is on the basis of theoretical expectations from the tests. It may be noted that a single MoP may be evaluated on the basis of multiple tests, in which case the average scores of all the tests corresponding to the MoP is used in EMMa. Readers interested in the detailed explanation results of these tests along with the motivation for scoring are referred to Appendix-A. For scoring *Usability*, the modeler is provided with the scoring system in Table 4 and the experience obtained after conducting the theorical tests to provide the input scores. Therefore, it becomes essential for the modeler to perform the tests to evaluate the MoPs, some of which may have a direct link while others are based on the experience of modeler, after performing the tests. MoPs belonging to Model Robustness and Computational Complexity have been obtained by running tests on a real-world network in Leuven, Belgium. Results from this test will be illustrated in the following section.

4.2. Scalability Test – Leuven Network

Most of the MoP's in EMMa are evaluated through theoretical test networks. The successful application of such networks lies in their ability to isolate specific model features. However, in practical applications transport models are used in real-world traffic networks. Therefore, it is imperative to understand and test the model behavior in such a real-world setting. Two MoP categories in EMMa are used to evaluate DTA models with these large-scale networks – *Model Robustness and Computational Efficiency*.

The network selected for this study is in the city of Leuven in Belgium, provided by Mint NV (Figure 16). The network consists of 430 centroids, 2697 links, 1832 nodes. The demand matrix originally provided was for the evening peak, starting from 15:00 to 20:00, disaggregated for every 15min (20 OD Matrices, excluding those for cooldown). The demand composed of car driver + freight (light and heavy). For the purpose of the current test, the demand was reduced to 17:00 (2hrs of demand) and the OD matrices were aggregated to 60mins (2 OD Matrices, excluding those for cooldown). This aggregation is performed to improve the speed of model run. As a comparative study is performed between the models, this simplification may be justified.

As the objective of the scalability test to analyze the model behavior under similar conditions, it is important to note their differences while performing the simulations. The route choice model in MARPLE uses C-Logit (default), whereas the StreamLine models were run using MNL. Furthermore, the time step of propagation for MARPLE and StreamLine: eGLTM was set to 1 sec, owing to their quick computation capabilities. For StreamLine: MaDAM, the propagation timestep was set to 2 sec, to improve the simulation speed. All the models were run for 35 iterations, which was used as stop criterion for convergence.

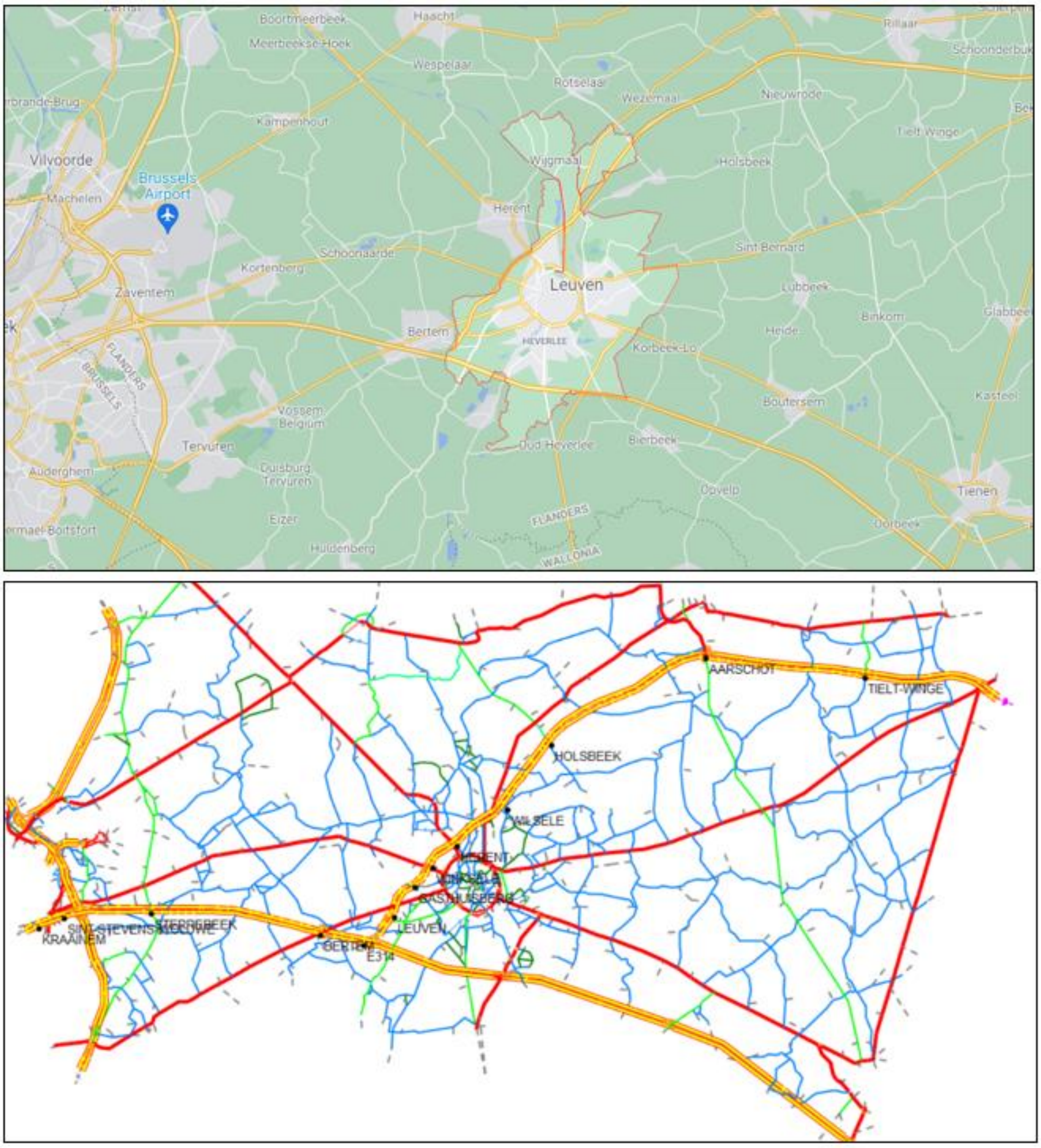


Figure 16: Leuven region in Belgium (Up) (Source: Google Maps), Leuven traffic network (Down)

The results from the scalability test on the three models are as shown in Table 7. The values are used as a direct input for evaluation in EMMa, as they are quantitative MoPs. It may be noted that the convergence error comparison is as described in Equation (10), Section 2.3.1.2 for MARPLE and Equation (12), Section 3.2.2 for StreamLine Models. It can be generalized that MARPLE performs exceedingly well in the computational efficiency MoP, owing to its quick run and low memory requirement.

Table 7: Results of scalability test - Quantitative MoP's in EMMa

Sl No	**MoP Name**	**MARPLE**	**StreamLine: MaDAM**	**StreamLine: eGLTM**
1	Convergence Error for final iteration in %	0.02%	7.86%	0.09%
2	Simulation Run Time in Hrs	2.1	185.8	2.6
3	Peak Memory Usage in MB	712.78	1768	3992

When compared to MaDAM, eGLTM performs substantially better in terms of simulation speed, although at the cost of higher memory requirement. Furthermore, eGLTM showcases a more stable state for equilibrium when compared to MaDAM due to a lower value for convergence error in gap function. This is further inspected through absolute duality gap values scattered over iteration number and calculation as shown respectively in Figure 17 and Figure 18.

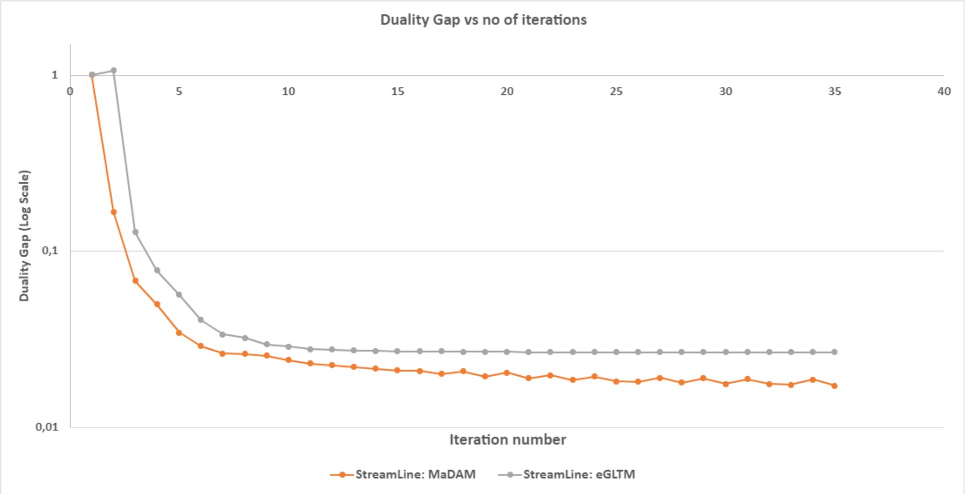


Figure 17: Duality gap vs No of iterations – StreamLine Models

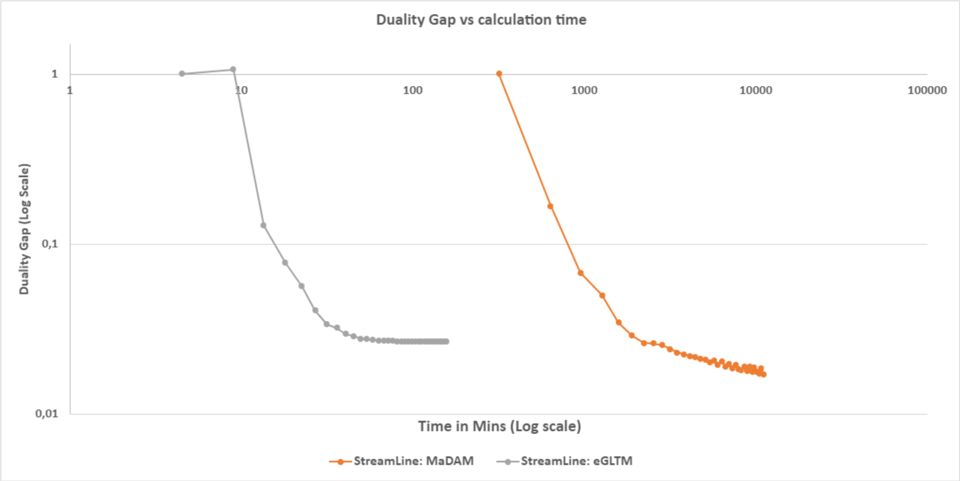


Figure 18: Duality gap vs Calculation time – StreamLine Models

It can be observed from Figure 17 that the convergence of MaDAM is slightly better than eGLTM at the 35th Iteration. However, this is possible at the expense of time (calculation time of MaDAM is almost 70 times of eGLTM) (Figure 18). The efficiency in propagation showcased by eGLTM makes its computation extremely quick than MaDAM, which was as per expectation. Although the duality gap value by itself is lesser for MaDAM, over the iterations, eGLTM performs better in terms of stability. However, we observe a strange value (greater than 1) for the duality gap for eGLTM, for iteration 2. This could mean that the route costs at this iteration is extremely high, compared to the cheapest route yielding a value for numerator in Equation (9). Further tests need to be performed to fully understand this anomaly.

While comparing the values of convergence error for the three models (Figure 19), both MARPLE and StreamLine: eGLTM showcases lower values of convergence error, indicating a more stable state of equilibrium. The convergence error is smoother in MARPLE as the values here are based on the relative change in flows as per Equation (10). The error values are also highly unstable in case of MaDAM.

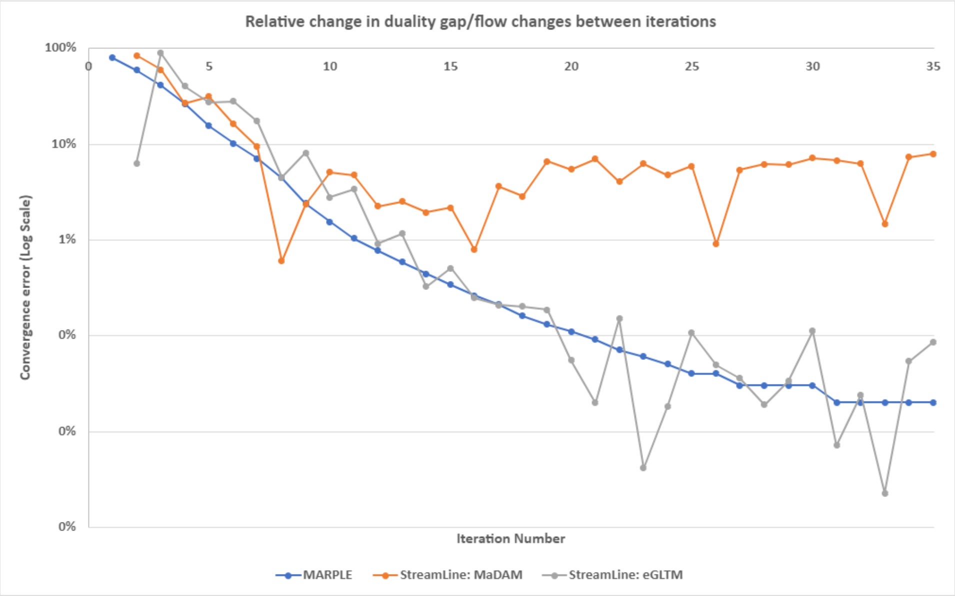


Figure 19: Convergence error vs Iteration Number – All three models

4.3. Results of Evaluation

The following section discusses in detail the results of the evaluation scores and the final ranking of the models. The final score per Macro DTA model is obtained after summation of the average MoP score for each of the 7 MoP categories. Note that the averaging that the scores provided in each of the MoP category are evaluated objectively and the number of MoPs in a particular category does not cause bias during the final ranking.

For **Strategic Planning** application, Figure 20 shows that across all model users, MARPLE scores better than the other two models. This is closely followed by a second rank for eGLTM by all model users. Upon detailed inspection of the MoP Category-wise scoring (example shown in Figure 21), It can be understood that the added edge for MARPLE was mainly in *Model Robustness.*

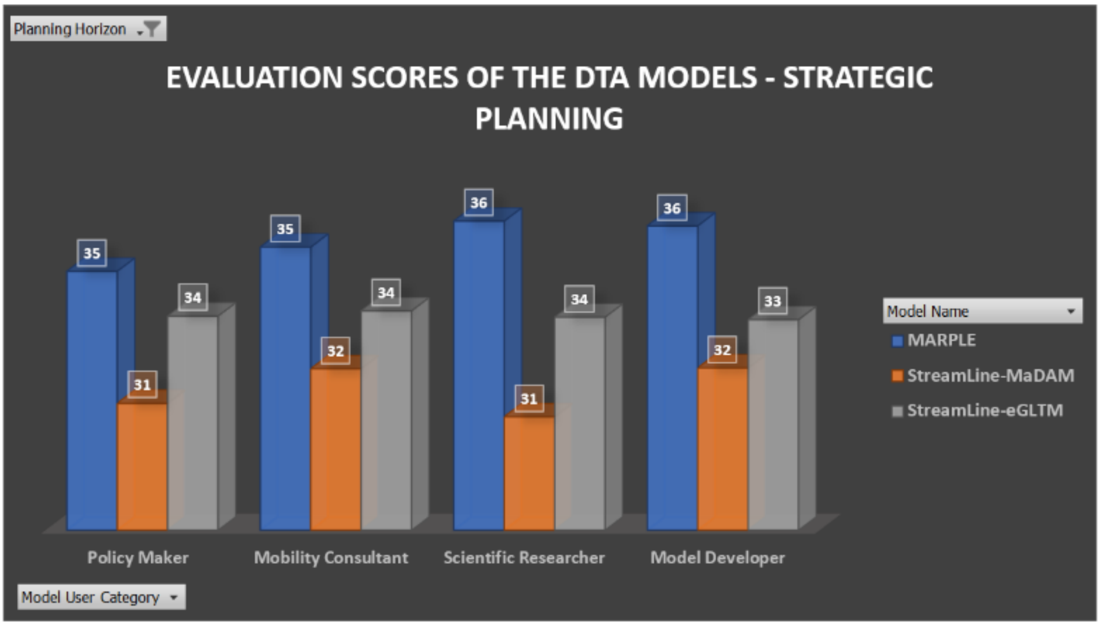


Figure 20: Final Scores of EMMa - Strategic Planning

The ability of MARPLE to converge faster to a stable equilibrium state compared to the other two models gave it a clear advantage, considering the importance given for *Model Robustness* by almost all model users. Ability of StreamLine: eGLTM to achieve a stable convergence has boosted its overall ranking closer to that of MARPLE.

As expected, the evaluation scores were almost similar across the three DTAs for the MoP categories of *Conceptual Validation*, which can be expected for a Strategic Planning application, as it was weighed with lesser importance by almost all model users. The added advantage of second order effects of StreamLine: MaDAM, portrayed through MoPs such as *Integration of Network Hierarchies*, was overshadowed by its poor performance in *computational efficiency* and *Model Robustness* MoPs. This is also the reason why StreamLine: eGLTM performed better than MaDAM.

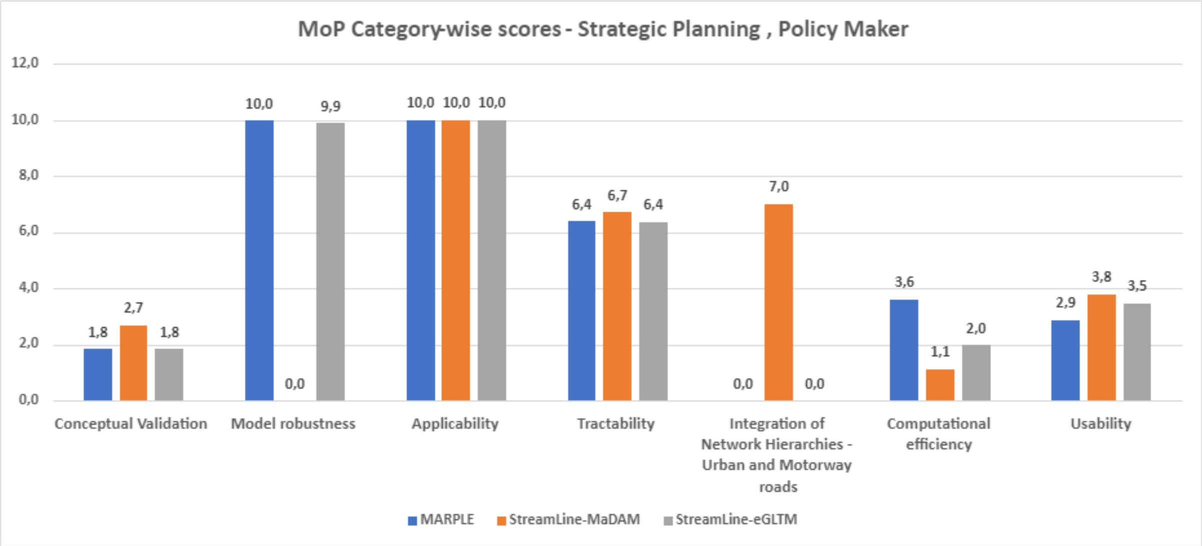


Figure 21: MoP category-wise scoring_Strategic Planning_Policy Maker

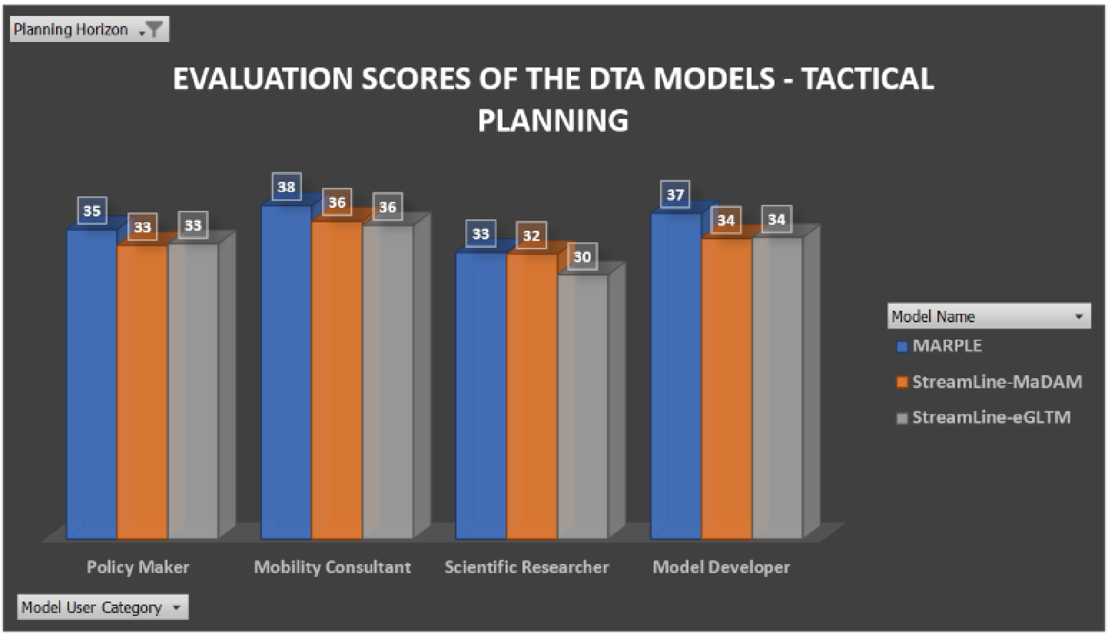


Figure 22: Final Scores of EMMa - Tactical Planning

For **Tactical Planning**, MARPLE scores better than the other DTA's, among most model users. However, the variation of final scores between the models are lesser compared to Strategic Planning

(Figure 22). This change in trend is further studied through detailed inspection as shown in Figure 23 and Figure 24.

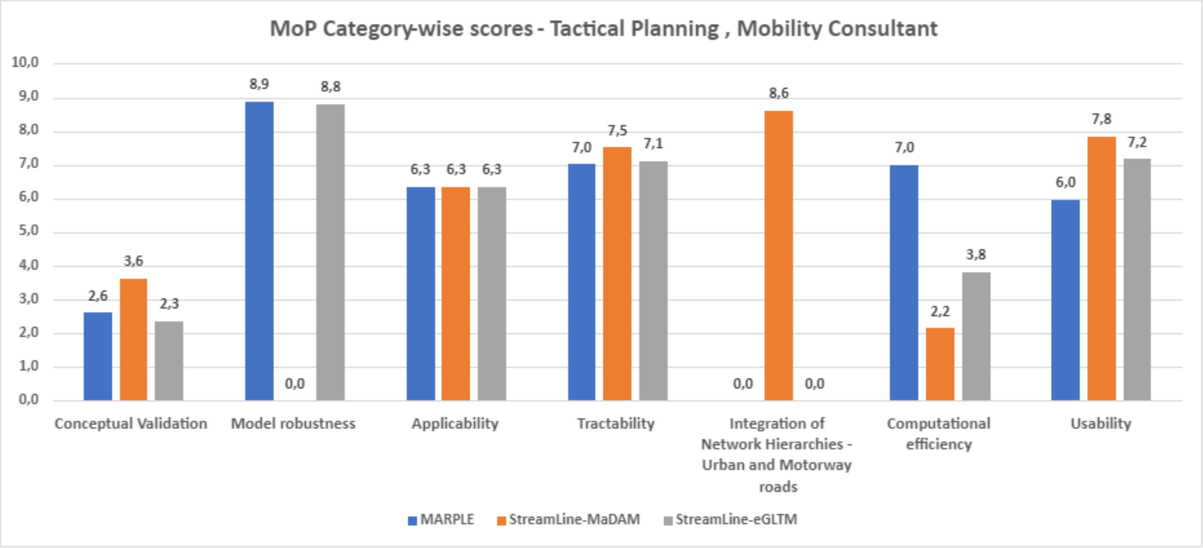


Figure 23: MoP category-wise scoring_Tactical Planning_Mobility Consultant

In case of a <u>Mobility consultant</u>(Figure 23), the larger importance associated with *Usability* and *Integration of Network hierarchies*, for Tactical Planning has boosted the overall scoring for MaDAM compared to Strategic Planning. The improvement of the score in these categories has managed to settle up for the zero score of MaDAM in *Model Robustness*. This has resulted in both the StreamLine models to have an almost equal score for Tactical Planning, which is the case amongst most model users.

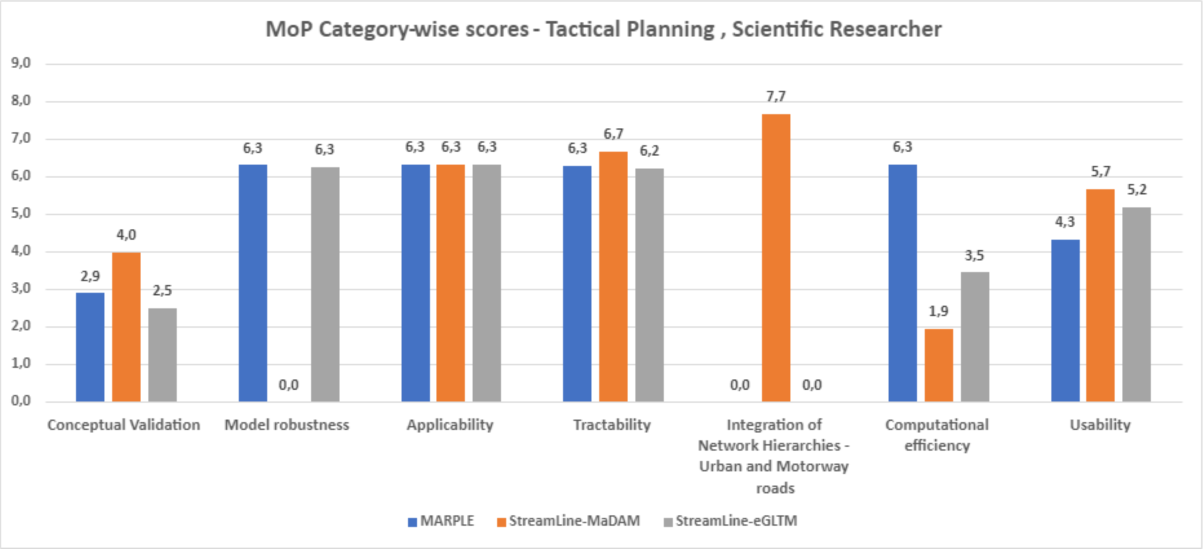


Figure 24: MoP category-wise scoring_Tactical Planning_Scientific Researcher

In case of a <u>Scientific Researcher</u> (Figure 24), the reduction in weight for *Model Robustness* has improved the score for MaDAM to such an extent that it ranked one of the highest among the three models. The same reason has caused the reduction of score for eGLTM, pushing it to the 3rd Rank. The advantage of MARPLE in *Computational efficiency* MoP has secured its spot among the top ranks, even though a smaller weight was associated with this MoP by the <u>Scientific researcher</u>.

For **Operational Planning**, both MARPLE and StreamLine: MaDAM secures the top rank amongst most Model Users. The score for eGLTM is slightly lower than the other models (Figure 25).

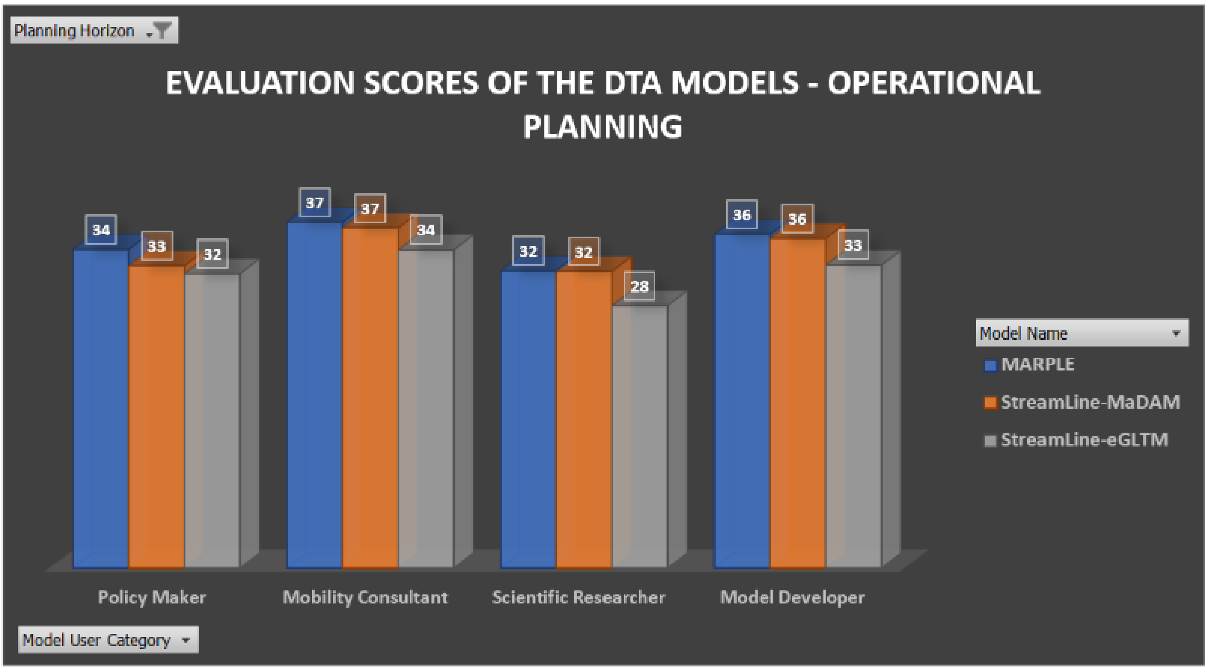


Figure 25: Final Scores of EMMa - Operational Planning

Upon further inspection at the model user level – Model Developer, it is observed that the higher weights associated with *Tractability*, *Integration of Network Hierarchies and Conceptual Validation* MoPs have countered the disadvantage of MaDAM in *Computational Efficiency*. The importance of the propagation model in Operational Planning is further manifested by the high weights associated with the above-mentioned MoPs. This importance has further boosted the overall score for MaDAM, when compared to both MARPLE and eGLTM. Even though eGLTM scored at par with MARPLE across most MoP categories and sometimes even better (*Usability*), an extremely fast simulation run with a very low memory requirement by MARPLE resulted in outranking eGLTM.

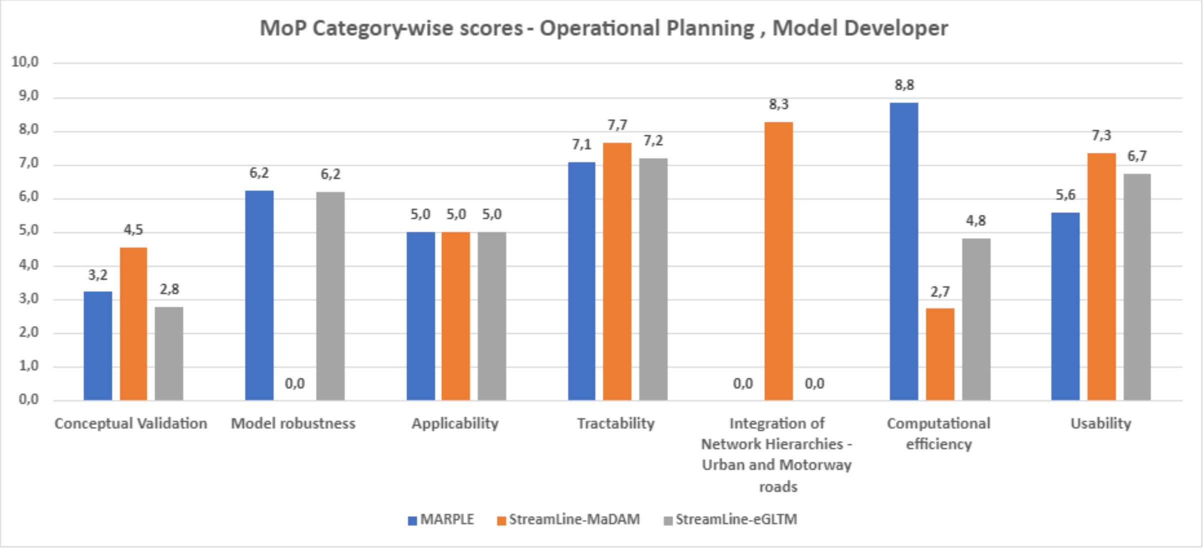


Figure 26: MoP category-wise scoring_Operational Planning_Model Developer

Results of EMMa can also be used to analyze the advantages and disadvantages of the DTA models, within each MoP category as shown through examples in Figure 27 and Figure 28. For *Tractability* – operational planning, scientific researcher perspective, it can be seen that StreamLine: eGLTM scores almost better than MARPLE except for the MoP which considers network control mechanism (*Propagation-Effect of link-level traffic controls*). Interestingly, the test results (Test 1.5.1 and 1.6.1, Appendix-A) shows that results were quite uncontrollable for StreamLine: MaDAM compared to MARPLE, which gave the latter a higher score for this MoP. The feature of traffic controls is still under development in the current version of eGLTM. Once this feature is enabled, the model will be superior in *Tractability* especially due to its theoretical accuracy and edge over StreamLine: MaDAM in *computational complexity*. The high scores for eGLTM in propagation -related MoPs are proof of this inference.

StreamLine: eGLTM is proved to be a good trade-off between MARPLE and MaDAM, as it provides the main advantages both the other DTA models, especially in case of an Operational Planning application. As expected, the scores of eGLTM are better than the others for the node model behaviors. This is because there exists an explicit node model to the link model in eGLTM, unlike MaDAM and MARPLE.

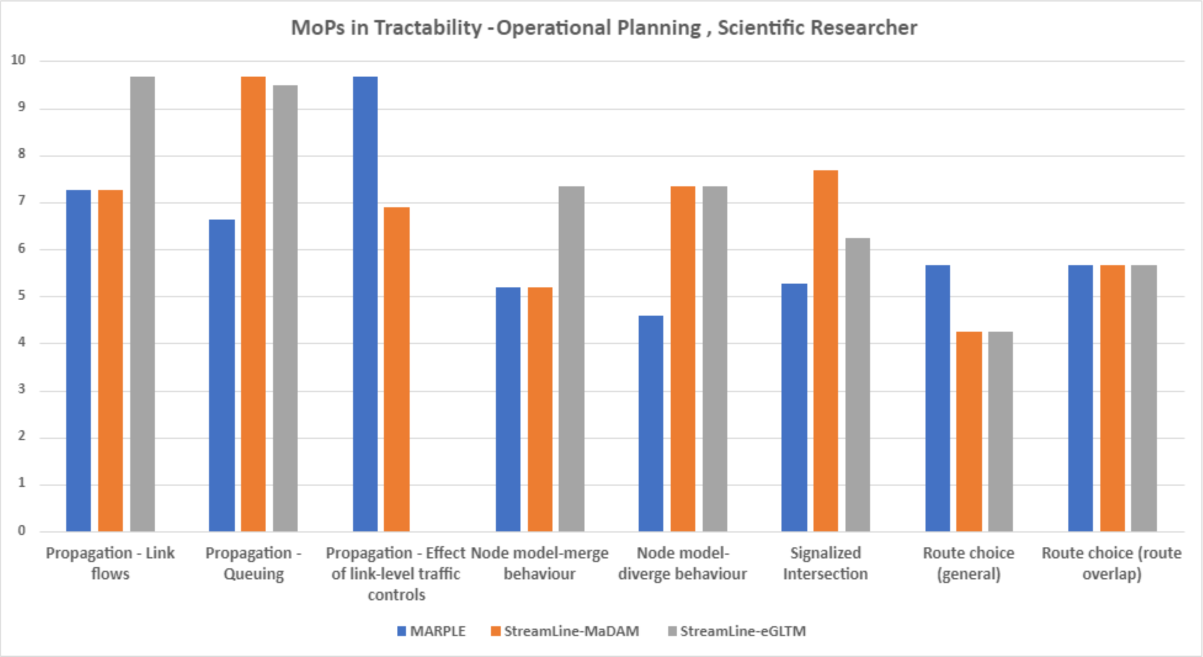


Figure 27: Scores for MoPs in Tractability, Operational Planning, Scientific Researcher Perspective

For *Conceptual Validation* – Tactical Planning, Model Developer perspective, it can be observed, that MARPLE gets a competitive score with StreamLine: MaDAM, mainly due to the smoothness of the link flow distribution. This can be attributed to the propagation and link speed calculations based on travel time functions as described in Section 2.3.1.3.

The main advantage for both the StreamLine models in this MoP category is the accurate identification and modeling of queues and its subsequent spillback to the links upstream. MARPLE, has scored relatively lesser here, due to its issues with dealing short links as covered in Test 1.3.1 in Appendix-A. *Conceptual validation* was given a high weightage by a model developer for tactical planning application (Figure 11).

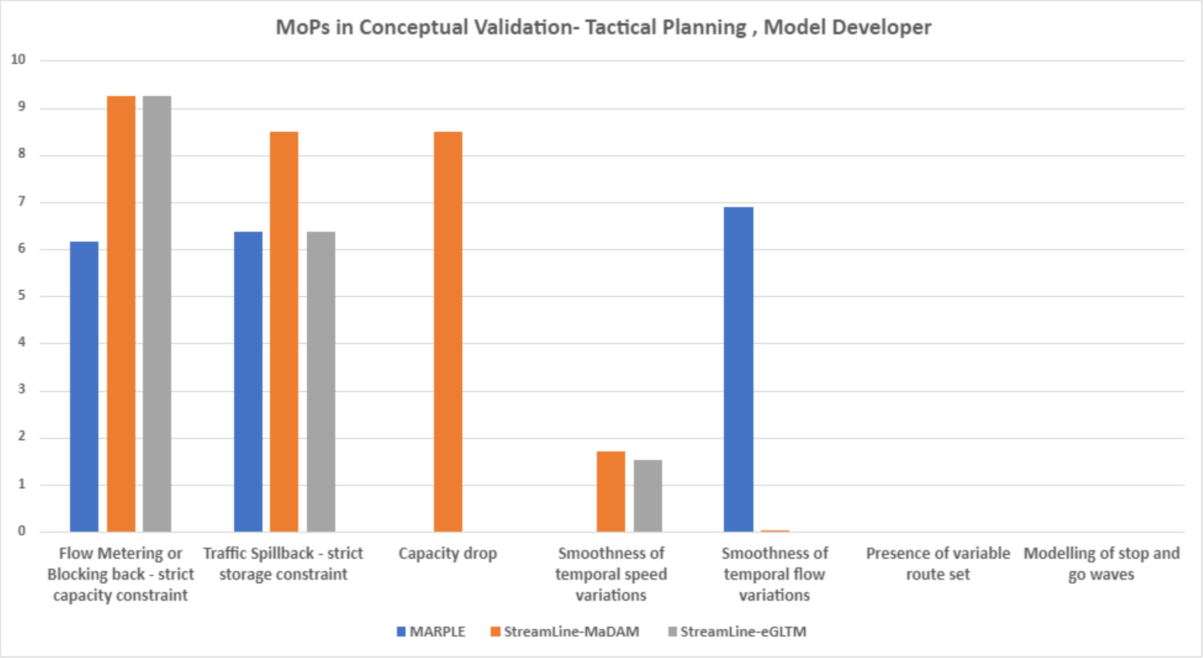


Figure 28: Scores for MoPs in Conceptual Validation, Tactical Planning, Model Developer Perspective

Variations in the overall score incurred in each model can be closely studied using EMMa, for evaluating the strengths and weaknesses of the model specific to an Application domain and/or Model User perspective. An example for the StreamLine: MaDAM is shown in Figure 29. The sensitivity is analyzed here with the base case as Strategic Planning. It can be observed, that amongst all model users, the overall scoring of MaDAM improves substantially as the application horizons gets smaller and the network spatially finer. The improvement is highest for a mobility consultant with 12% and 14% for Tactical and Operational respectively. As already described the strengths of MaDAM in the MoP categories of *Integration of Network hierarchies*, *Tractability* and *Usability* have boosted its score throughout the planning horizons, especially since these MoPs were given high weightage by the Mobility Consultant. For a Policy maker, larger improvement in scores is observed for Tactical (6%) when compared to Operational Planning (4%). The smaller weight allotted for *Applicability* in Operational Planning by the Policy Maker has reduced the overall scoring for MaDAM. Similar sensitivity studies can be studied for each model per application domain and model user perspective.

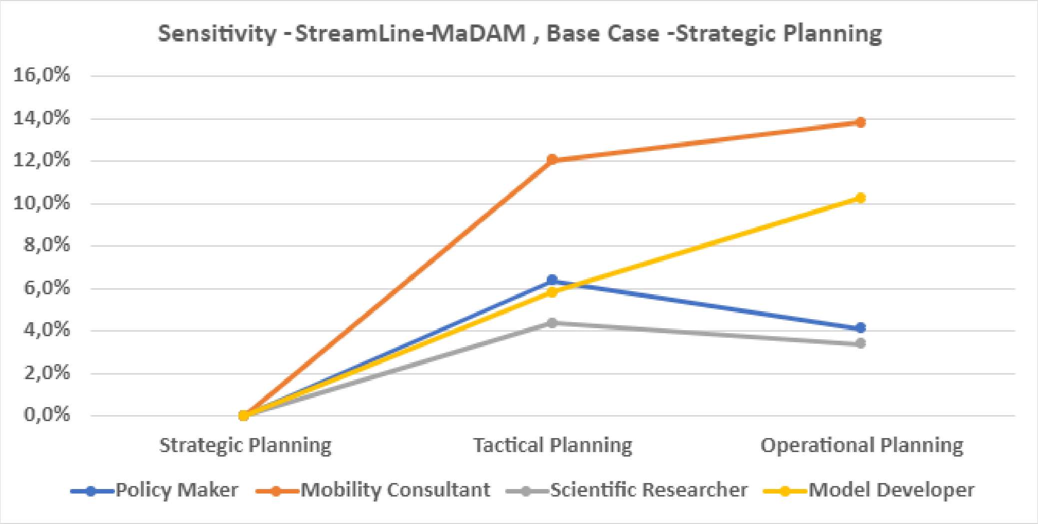


Figure 29: Sensitivity Analysis - StreamLine: MaDAM, Base Case Strategic Planning

5. Conclusions and Future Recommendations

The primary aim of this research was to develop a framework for evaluating Macroscopic DTA models. As described in the preceding sections, this involved the formulation of certain measures which can be yardsticks for comparing the strengths and weaknesses of the models. The yardsticks or MoPs, ensure that the evaluation is performed objectively without holding a bias regarding the individual preference of a model against another. The evaluation as explained in the methodology is motivated heavily from Multi Actor Multi Criteria Decision Analysis (MAMCDA). The evaluation of the alternatives or the DTA models, are comparative. The validation of EMMa is conducted by applying the framework for evaluating three Macroscopic DTA models commonly used in the Netherlands. In the light of this research context, the final chapter of this thesis report is divided into three; The first part will describe the general conclusions from the application of EMMa (Section 5.1). The second section will describe the answers to the research questions formulated in the project proposal phase (Section 5.2), and the third section will discuss the drawbacks and future recommendations for EMMa (Section 5.3).

5.1. Key Takeaways from the Results

The conclusions and inferences from the results of EMMa are as summarized below

- For **Strategic Planning**, the most important modeling property for a Macroscopic DTA amongst all model users was *Model Robustness*. The achievement of a stable and accurate equilibrium state was given one of the highest priorities while considering the application of a DTA. The underlying factor considered here was also the size and complexity of the traffic network, as for strategic planning application, the expectation would be a large-scale network.

- For **Tactical Planning**, the importance of *Model Robustness* remained high across all model users, except for a Scientific Researcher. A variability of weights could be observed amongst the results, where experts in the field of traffic modeling such as a Scientific Researcher and a Model Developer felt the need for a faster model, whereas a policy maker felt this of lesser importance and thus lesser weightage was associated with it.

- As the planning horizon became smaller, as is the case with **Operational Planning**, the importance associated with the finer details of model run such as the propagation behavior (*Tractability*), the smooth *integration of network hierarchies*, the modeling of real-world effects (*Conceptual Validation*) such as capacity drop, smooth variation of speeds and flows within the model etc becomes more important across all model user categories. The importance associated with *Model Robustness* was the least for Operational Planning, especially for expert traffic modelers because achievement of true equilibrium is not feasible and not a strict requirement here (especially considering SDUE). The weightage associated with *computational complexity* was one of the highest for this application horizon, keeping in mind the usage of the model for real-time applications which requires frequent model runs and quick results for decision making purposes.

- Interestingly, the model *Usability* was not an important criterion by most model users. Mobility Consultants, however felt the need to work with a model with exceptional ease of use, owing to their hands-on and practical interaction with traffic models on a regular basis, which may be added with a lack of detailed knowledge in traffic modeling unlike the experts in traffic studies.

- Looking at the scalability results between the StreamLine Models, eGLTM has a substantially faster simulation run, when compared to MaDAM. Furthermore, eGLTM showcases a more stable state for equilibrium due to a lower value for convergence error in gap function.

However, if the modeler has enough time to spare, StreamLine: MaDAM converges to a smaller value of relative dynamic duality gap, indicating a more accurate result (Figure 17 and Figure 18). The simulation run time of MARPLE is slightly better to StreamLine: eGLTM, with an achievement of more stable state of equilibrium. Thus, MARPLE exhibits a blazing fast simulation with a relatively stable convergence to equilibrium.

- The results of EMMa showed a variation in model rankings across different application horizons. For **Strategic Planning**, the results were in favor of MARPLE followed by StreamLine: eGLTM across all model user categories. Upon closer inspection, the advantage of these DTA's was in the MoP category - *Model Robustness*. This was clearly evident for a Policy Maker and a Mobility consultant, who felt a greater need for a stable state of equilibrium, compared to other features of a DTA. However, for experts in DTA models such as Scientific Researchers and Model Developers required a faster model run even for strategic planning, which placed MARPLE over StreamLine: eGLTM for these users. Even though the computation times of both these DTA's were comparable, MARPLE simulations had smaller memory requirements. As per expectation, StreamLine: MaDAM secured the lowest rank across all model users, owing to its extremely slow computation and unstable convergence to equilibrium. The strength of MaDAM lies in its ability to mimic the propagation behavior accurately which was given not a priority in Strategic Planning.

- In Case of **Tactical Planning**, MARPLE performed slightly better than the StreamLine models across all model user categories. As expected, the overall score for StreamLine: MaDAM improved substantially, when compared to Strategic Planning owing to the larger importance associated with accuracy of propagation model (MoP category – *Tractability*). The ability of the second-order CTM in MaDAM to seamlessly integrate urban and non-urban links gave an added advantage as it was an important criterion for evaluation especially among Policy Makers and Mobility Consultants. However, this complexity as at the expense of a larger simulation run time, which made both StreamLine models rank equally across all model user categories.

- As the spatial granularity got smaller as is the case for **Operational Planning** application, both MARPLE and StreamLine: MaDAM secured similar scores across all model user categories. The model users weighed the quality of the model results, theoretical tractability (StreamLine: MaDAM), accuracy of the propagation model (StreamLine: MaDAM), Integration of Network Hierarchies (StreamLine: MaDAM) and computational complexity (MARPLE) to be very high for smaller application horizons. For detailed analysis at link-level as is the case with most Operational Planning applications StreamLine: MaDAM should be the preferred choice, at the expense of some additional simulation time. With improvements in incorporating traffic controls (Prototype of which is under development currently) and additional bug fixes in propagation model, the theoretical tractability of StreamLine: eGLTM, will make it superior to both MARPLE and StreamLine: MaDAM, enabling a quick model run without compromising much on link-level *Tractability*. This is especially true for quick-scan strategic planning applications and real-time operational planning applications.

5.2. Answers to Research Questions

This subsection would summarize the answers to the main research question and the sub-questions which were answered in detailed in the preceding sections of this report. The answers to the sub-questions are provided first which helps in answering the primary research question.

Sub Question-1: How representative are the DTA models chosen for conducting the current research? what are their strengths and weaknesses? (Qualitative classification)

As stated, before the DTA models selected for this case study are on the basis of a qualitative classification scheme as stated in Section 2. As propagation or dynamic network loading sub-module of the DTA is one of the main elements of focus in this research, the models selected differs from each other in this aspect. The propagation model in MARPLE is on the basis of link-performance functions, which differs from link to link on the basis of its link type (Controlled/Normal). The DTA models under the StreamLine framework MaDAM and eGLTM are essentially based on Traffic flow theory (i.e. fundamental diagrams) and belong to second-order and first order models respectively on the basis of the number of traffic variables.

Between the StreamLine models, MaDAM is based on the Cell Transmission Model using the Van Aerde fundamental diagram and eGLTM is an event-based algorithm for the link transmission model which can use any concave fundamental diagram (but in this study, the quadratic-linear diagram was used). Essentially, eGLTM removes temporal discretization errors from the regular LTM, whereas the regular LTM had already removed spatial discretization errors from the CTM. On top of that, the removal of space and time discretization makes eGLTM computationally efficient, which is also experimentally proven with this research project.

All three models used in this research project are capacity and storage constrained models and features "blocking back" in the bottleneck links to create horizontal queuing onto the links upstream. These classifications motivate the choice of the DTA models for the research. Furthermore, the availability to software and guidance along with the accessibility to the models played a major role in the choice of these macroscopic DTAs.

Sub Question-2: What are the measures of performances that will be used to evaluate under each application scenarios – strategic/tactical/operational, and with different model user perspectives?

MoPs form the primary yardstick for the objective evaluation in this research. Therefore, the MoPs selected for this research are motivated on the basis of numerous literature sources and discussions with traffic modeling experts as mentioned in Section 2.4. On the basis of the scope defined for this project, 28 MoPs are identified which are classified in a two layered system as mention in Section 3.2. The primary classification involved the categorization into 7 main titles: *Conceptual Validation*, *Model robustness*, *Applicability*, *Tractability*, *Integration of Network Hierarchies - Urban and Motorway roads*, *Computational efficiency* and *Usability*. The secondary classification involved the measurement type of the MoP – Qualitative, Quantitative and Binary measurements.

Sub Question-3: How to score and rank the models on the basis of the evaluation criterion?

After the identification of the MoPs, the DTA's are scored and ranked. The majority of the MoPs identified for the evaluation are qualitative, mainly because empirical validation of the models lies outside the scope of this research. Each of the 28 MoPs are inspected and evaluated by performing simulation tests. As a single model component is separated and its behavior is studied, simple theoretical test networks are successfully employed to serve this purpose. This led to the formation of 32 tests which helped in scoring majority of the MoPs. A large-scale real-world network in Leaven, Belgium is also used to evaluate MoPs related to scalability such as *Model Robustness* and *Computational Complexity*.

A qualitative score system is created for evaluation, which has an aggregated interval keeping in mind the sensitivity it may have due to uncertainty. The qualitative scoring was based on the difference between the expectation from a model behavior given the underlying (traffic flow or mathematical) theory vs the actual behavior as seen through the simulation results. Certain other MoPs, such as *Usability*, had a different score system as they were representing the ease of use of the model. The scoring relied little on the expertise of the modeler, but rather was dependent on the experience he/she develops after performing the tests. The testing series and the networks used for the same are as described in Table 3 and Appendix-A. Most of the tests are motivated from literature, interviews with experts and the expertise of the researcher.

On the basis of the evaluation, the MoP scored are then normalized to make the comparison Apple-to-Apple. The next step is to average the MoP scores, calculated per MoP category as described in Section 3.5. The normalized scores are then multiplied with the weights obtained from the Model User survey (using questionnaires and interviews with experts) which provides the added dimensions of the model user category and application horizon. The average scores per MoP category is then summed to provide the final scores per DTA Model, which provides the model ranking.

Primary Research Question: How to compare Macroscopic Dynamic Traffic Assignment Models based on their performance under various evaluation themes?

The comparison of the DTA's have been performed in two layers – objective evaluation and subjective evaluation. The objective side evaluates the models purely on the basis of simulation results. However, the research also entails upon the subjective dimension to the evaluation. The subjective side showcases the differences in importance associated with model features which varies from model users to application domains.

Departing from the selection of the DTA's, the evaluation themes used for comparing the models is delegated in two dimensions – **Primary** (MoPs – Section 3.2) and **Secondary** (Application domain and Model User perspective – Section 3.3). The model performance is evaluated using the MoPs, which is scored using a series of testing on theoretical networks (Appendix-A) and a real-world large-scale network (Section 4.2). This step results in the initial score matrix (Section 4.1). The comparison of the MoP scores is then made apple-to-apple by means of standardization (Section 3.5), which results in the final evaluation scores of the models (Section 4.3). The weights provided by the model user survey incorporates the subjective side of evaluation. Thus, the design, development and validation of the evaluation framework – EMMa, is the answer to the main research question. The framework also acts as an experimental comparison of the strengths and weaknesses of the Macroscopic DTA models under comparison, re-validating their theoretical descriptions.

5.3. Limitations and Future Recommendations

The formulation of any model or framework comes with its scope for improvement, which adds another dimension in its development cycle. Similarly, the framework developed as part of the current research - EMMa has ample scope for improvements which can be taken up as an extension for future works. Some of them are as listed below:

- **Limitation** – One of the properties of a DTA model which validates its application in actual case studies is accuracy. This property refers to the accuracy of the simulated results to that of empirical data which is usually tested through statistical relationships of closeness such as coefficient of determination: R-squared value, standard error etc. However, due to lack of empirical data and availability of time, this property was left out from the current research and would have added an important dimension for an MoP in EMMa.

- ➢ **Future Research direction** – The test for this MoP could be performed over an actual case network at a link level by comparing the results simulated by the model to that of empirical data of the same link or road section in the actual network. Measures such as speed distributions, merge-diverge behavior etc. could be tested, with statistical quantitative MoPs describing the fitness of the simulated results. The tests can also be an extension to the empirical validation of the model under analysis.

- **Limitation** – MoPs related to multi-class modeling is limited to binary measurements for the current version of EMMa.

- **Future Research direction** – However, for future versions, a qualitative MoP could be included for evaluating the multi-class modeling behavior. On the basis of availability of empirical data, this MoP could also be quantitative describing the closeness of the results.

- **Limitation** – The survey performed as part of this research for obtaining the weights, were limited to 15 responses. This was mainly because the research was part of a thesis project and there was a limitation of time.

- **Future Research direction** – Literature recommends surveying close to 100 respondents spread across various model user categories, possibly from different countries to obtain impartial and unbiased weights. Moreover, the current research takes multiple responses from a single respondent, as he/she belongs to more than a single model user category. However, this can create a bias in the responses. While extending the model user survey, care should be taken that a single respondent should be providing answers to a single model user category to make the weights unbiased.

- **Limitation** – The normalization technique adopted for the current version of EMMa is Linear: Max. For comparison purposes other techniques as mentioned in Table 5, were explored and the sensitivity in the results were identified. Linear: Max method indicated the maximum variability in the results amongst the models. The method was also successful in normalizing the nil value scores for certain MoPs, within EMMa. However, Linear: Max is a rather strict normalization technique (as observed through the *Model Robustness* score for StreamLine: MaDAM), which may not be desired.

- **Future Research direction** – In future versions of EMMa, additional normalization techniques may be explored and the modeler can be provided with the best normalization technique suitable for his/her evaluation.

- **Limitation** – As mentioned in the Methodology (Section 3.2.2), the modified MoP for Model Robustness, used for the current application of the three DTA's is based on relative change in gap function between iterations (Equation (12)). This modification was performed for the current application because convergence in MARPLE was based on relative change in flow values between iterations (Equation (10)). Although the convergence error values from both these measures are comparable, this is not strictly accurate. Moreover, this workaround tests the stability of the equilibrium state which is translated to a MoP for model robustness.

- **Future Research direction** – For future comparison, the absolute value of Dynamic duality gap at the end of a fixed number of iterations, should be the criterion to decide the model robustness (Equation (9)). This MoP evaluates the accuracy of equilibrium state.

- **Future Research direction** – Extending further on the tests for scalability, actual case networks can be tested for both tactical and operational planning applications. However, MoPs that reflect the evaluation based on these application horizons need to identified to be included in EMMa.

- **Future Research direction** – As there is minimal literature on frameworks for evaluating DTA models, EMMa could also be extended to include microscopic, mesoscopic and Hybrid DTA models. However, this will include larger number of MoPs with model user surveys extending over 200-300 respondents (roughly) as the evaluation criteria will be plenty.

- **Future Research direction** – The scope for the current research was restricted to the time horizon in DTA application. The evaluation framework could also be extended to include other types of model application, such as short-term forecasting, optimization, impact assessment, online/offline applications etc.

The ability of a DTA model or any transport model for that matter, is to simulate the behavior of a transport system within a virtual environment, which acts as a safe haven for trails and experiments. It becomes clear to any model user or a developer that an ideal model does not exist but rather serves as a tool for decision-making for the problem at hand on the basis of some theoretical assumptions. Thus, the choice of the model is a key criterion in finding solutions to the problem. The framework EMMa thus serves as a model for macroscopic DTA models to help the modeler to choose the right model. The additional dimensions of the framework provide the various perspectives with which the model can be used. From the application of EMMa to the three models selected for this research, the fundamental trade-off between model complexity and computational speed was clearly visible from the results.

MARPLE owing to high-speed computation capabilities and faster achievement of a stable equilibrium state proved to achieve Rank No.1 across most model user categories and application horizons. This can be interpreted that, model users in general prefer these characteristics over complexity of results (through various complex features of the model as is the case with MaDAM). However, we observe variations across model users for model preference, which validates our original hypothesis that the right choice of a model primary depends on the person using it and the application it is deployed for. Inclusion of a larger spectrum of model user surveys might alter these weights, but that is subject to future work.

6. Bibliography

Aimsun. (2021, September 10). *Aimsun Next*. Retrieved from www.aimsun.com: https://www.aimsun.com/aimsun-next/

Akçelik, R. (1991). Travel time functions for transport planning purposes: Davidson's function, its time dependent form and alternative travel time function. *Australian Road Research 21*, 49-59.

Akçelik, R. (2003). *Speed-Flow Models for Uninterrupted Traffic Facilities.* Melbourne, Australia: Technical Report, Akcelik and Associates Pty Ltd.

Aw, A., & Rascle, M. (2000). Resurrection of "second order" models of traffic flow. *SIAM J Applied Mathematics 69*, 916-938.

Bar-Gera, H. (2010). Traffic assignment by paired alternative segments. *Transportation Research Part B Methodology 44*, 1022-1046.

Ben-Akiva, M., B., Koutsopoulos, H., & Mishalani, R. (1998). DynaMIT: A Simulation-basedSystem for Traffic Prediction and Guidance Generation. *3rd Triennial Symposium on Transportation Systems.* San Juan, Puerto Rico.

Binsbergen, A. v. (2020). Lecture on Multi Criteria (Decision) Analysis. *CIE5817-Assessment of transport infrastructure and systems:*. TU Delft.

Bliemer, M. (2001). *Analytical Dynamic Traffic Assignment with Interacting User-Classes: Theoretical Advances and Applications Using a Variational Inequality Approach.* Delft, The Netherlands: PhD Thesis, Delft University of Technology.

Bliemer, M. C., & Raadsen, M. P. (2019). Continuous-time general link transmission model with simplified fanning, Part I: Theory and link model formulation. *Transportation Research Part B*, 442-470. doi:https://doi.org/10.1016/j.trb.2018.01.001

Bliemer, M. C., Raadsen, M. P., Brederode, L. J., Bell, M. G., Wismans, L. J., & Smith, M. J. (2017). Genetics of traffic assignment models for strategic transport planning. *TRANSPORT REVIEWS, 37*(1), 56-78. Retrieved from http://dx.doi.org/10.1080/01441647.2016.1207211

Bliemer, M., & Bovy, P. (2008). Impact of Route Choice Set on Route Choice Probabilities. *Transportation Research Record 2076*, 10-19. doi:10.3141/2076-02

Bliemer, M., & Taale, H. (2006). Assignement, Route Generation and Dynamic Traffic Assignment for Large Networks. *First International Symposium on Dynamic Traffic Assignment*, (pp. 90-99). Leeds, United Kingdom.

Bliemer, M., Raadsen, M., Smits, E.-S., & Romph, E. d. (2013). Requirements for traffic assignment models for strategic transport planning: A critical assessment. *Working Paper ITLS-WP-13-16*, 1-28.

Bliemer, M., Versteegt, H., & Castenmiller, R. (2004). INDY: A New Analytical Multiclass Dynamic Traffic Assignment Model. *TRISTAN V Conference.* Guadeloupe.

Boyce, D., Ralevic-Dekic, B., & Bar-Gera, H. (2004). Convergence of traffic assignments: how much is enough? *Journal of Transportation Engineering of ASCE 130*, 49-55.

Brederode, L., Bliemer, M., & Wismans, L. (2010). STAQ: Static Traffic Assignment with Queuing. *European Transport Conference.* Glasgow, United Kingdom.

Brederode, L., Pel, A., Wismans, L., de Romph, E., & Hoogendoorn, S. (2019). Static Traffic Assignment with Queuing: model properties and applications. *Transportmetrica A: Transport Science 15:2*, 179-214. doi:https://doi.org/10.1080/23249935.2018.1453561

Bureau of Public Roads. (1964). *Traffic Assignment Manual.* Washington: Burueau of Public Roads (United States).

Calvert, S., Minderhoud, M., Taale, H., Wilmink, I., & Knoop, V. L. (2016). *Traffic assignment and simulation models: State-of-the-Art Background Document.* Delft, Netherlands: TrafficQuest, Centre for Expertise on Traffic Management.

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks. *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, (pp. 697-711). Lyon, France.

Chen, A., Kasikitwiwat, P., & Ji, Z. (2003). Solving the Overlapping Problem in Route Choice with Paired Combinatorial Logit Model. *Transportation Research Record Journal of the Transportation Research Board. 1857*, 65-74. doi:10.3141/1857-08.

Chen, H.-K. (1999). *Dynamic Travel Choice Models - A Variational Inequality Approach.* Heidelberg, Germany: Springer.

Chiu, Y.-C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, T., & Hicks, J. (2011). Dynamic Traffic Assignment: A Primer. *Transportation Research E-Circular*.

Daganzo, C. (1993). *The cell transmission model part I: a simple dynamic representation of highway traffic.* Berkeley: University of California.

Daganzo, C. (1994). The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B 28*, 269-287.

Daganzo, C. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B 29*, 79-93.

Dial, R. (2006). A path-based user-equilibrium traffic assignment algorithm that obviates. *Transportation Research Part B Methodology 40*, 917-936.

FakhraeiRoudsari, F., Huang, W., & Tampère, C. (2015). *Benchmarking StreamLine Dynamic Traffic Assignment model: theoretical test cases.* Diepenbeek, Belgium: Steunpunt Verkeersveiligheid, Consortium UHasselt, KU Leuven en VITO.

Ferrara, A., Sacone, & Siri, S. (2018). Freeway Traffic Modelling. In *Freeway Traffic Modelling and Control. Advances in Industrial Control.* Springer, Cham. Retrieved from https://doi.org/10.1007/978-3-319-75961-6_5

Flügel, S., Flötteröd, G., Kwong, C. K., & Steinsland, C. (2014). *Evaluation of methods for calculating traffic assignment and travel times in congested urban areas with strategic transport models.* Institute for Transport Economics, Norwegian Centre for Transport Research.

Gadelmawla, E., Koura, M., Maksoud, T., Elewa, I., & Soliman, H. (2002). Roughness parameters. *Journal of Materials Processing Technology.123*, 133-145. doi:10.1016/S0924-0136(02)00060-2.

Garavello, M., & Piccoli, B. (2016). *Traffic flow on networks.* American Institute of Mathematical Sciences.

Garavello, M., Han, K., & Piccoli, B. (2006). *Models for vehicular traffic on networks.* American Institute of Mathematical Science.

Gentile, G. (2010). Chapter 8 - The general link transmission model for dynamic network loading and a comparison with the DUE algorithm. In G. Gentile, *New Developments in Transport Planning: Advances in Dynamic Traffic Assignment.*

Gentile, G. (2011). Chapter 8 - The general link transmission model for dynamic network loading and a comparison with the DUE algorithm. In G. Gentile, *New Developments in Transport Planning: Advances in Dynamic Traffic Assignment.*

Godunov, S. (1959). A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik 47*, 271-306.

Heydecker, B., & Addison, J. (2005). Analysis of dynamic traffic equilibrium with departure time choice. *Transportation Science 39*, 39-57.

Hoogendoorn, S., & Bovy, P. (2001). State-of-the-art of vehicular traffic flow modeling. *Journal of Systems and Control Engineering 215*, 283-303.

Kessels, F., Lint, J., Vuik, C., & Hoogendoorn, S. (2014). Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics 4,* 1-29.

Knoop, V. L., Hegyi, A., Salomons, M., Lint, H. v., Yuan, Y., & Landman, R. (2019, November 8). Traffic Flow Modelling & Control - CIE4825 and CIE5821 Lecture notes. Delft, Zuid-Holland, Netherlands: Civil Engineering and Geosciences (CEG), Transport & Planning - Delft University of Technology.

Larsson, T., & Patriksson, M. (1995). An augmented lagrangean dual algorithm for link capacity side constrained traffic assignment problems. *Transp. Res. Part B Methodol*.

Lebacque, J. (1996). The Godunov scheme and what it means for first order traffic flow models. *Proceedings of the 13th international symposium on transportation and traffic theory*, (pp. 647-677).

Lighthill, M., & Whitham, G. (1955). On Kinematic waves II: A theory of traffic flow on long crowded roads. *roayal society of London, Part A 229 (1178)*, (pp. 281-345). London.

Mahmassani, H., Chiu, Y.-C., Chang, G., Peeta, S., & Ziliaskopoulos, A. (1998). *Off-line Laboratory TestResults for the Dynasmart-X Real-Time Dynamic Traffic Assignment System.* Texas: Technical Report ST067-85-TASK-G, Center for Transportation Research, The University of Texas at Austin.

Messmer, A., & Papageorgiou, M. (1990). METANET: a macroscopic simulation program for motorway networks. *Traffic Engineering and Control, 31*, 466-470. Retrieved from https://www.researchgate.net/publication/282285780

Nemhauser1, G. L., & Merchant, D. K. (1978). A Model and an Algorithm for the Dynamic Traffic Assignment Problems. *Transportation Science 3*, 183-199.

Nemhauser2, G. L., & Merchant, D. K. (1978). Optimality Conditions for a Dynamic Traffic Assignment Model. *Transportation Science 12*, 200-207.

Ni, D., Leonard, J. D., Guin, A., & Williams, B. M. (2004). Systematic Approach for Validating Traffic Simulation Models. *Transportation Research Record Journal of the Transportation Research Board 1876(1876)*, 20-31.

OmniTRANS Transport Planning Software. (2016). Traffic Assignment - Static and Dynamic (StreamLine). In O. T. Software, *OmniTRANS User Manual.*

Ortúzar, J. d., & Willumsen, L. (2011). *Modelling Transport - 4th Edition.* Oxford, United Kingdom: John Wiley and Sons, Ltd.

Pang, J., Han, L., Ramadurai, G., & Ukkusuri, S. (2012). A continuous-time linear complementarity system for dynamic user equilibria in single bottleneck traffic flows. . *Math. Program. 133*, 437-460.

Papageorgiou, M. (1990). Modelling and real-time control of traffic flow on the Southern part of Boulevard Périphérique in Paris: part I: modelling. *Transportation Research Part A 24*, 345-359.

Papageorgiou, M. (1998). Some remarks on macroscopic traffic flow modelling . *Transportation Research Part A 32*, 323-329.

Papageorgiou, M., Blosseville, J.-M., & Hadj-Salem, H. (1989). Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. *Transportation Research Part B 23*, 29-47.

Patil, P. N., Ross, K. C., & Boyles, S. D. (2021). Convergence behavior for traffic assignment characterization metrics. *Transportmetrica A: Transport Science, 17*(4), 1244-1271. doi:10.1080/23249935.2020.1857883

Payne, H. (1971). Models of freeway traffic and control. Mathematical models of public systems. *Simulation council proceedings, volume 1*, 51-61.

Peeta, S., & Ziliaskopoulos, A. K. (2001). Foundations of Dynamic Traffic Assignment: The past, the present and the Future. *Networks and Spatial Economics, 1*, 233-265.

Pravinvongvuth, P. S., & Chen, A. (2005). Adaption of the paired combinatorial logit model to the route choice problem. *Transportmetrica. 1.*, 223-240. doi:10.1080/18128600508685649

Raadsen, M. (2018). *Aggregation and decomposition methods in traffic assignment: towards consistent and efficient planning models in a multi-scale environment.* Sydnet: University of Sydney, Sydney Business School. Retrieved from https://ses.library.usyd.edu.au/bitstream/handle/2123/18186/raadsen_m_thesis.pdf?sequence=2&isAllowed=y

Raadsen, M., Bliemer, M., & Bell, M. (2016). An efficient and exact event-based algorithm for solving simplified first order dynamic network loading problems in continuous time. *Transportation Research Part B (92)*, 191-210.

Raadsen, M., Mein, H., Schilpzand, M., & Brandt, F. (2010). Implementation of a single dynamic traffic assignment model on mixed urban and highway transport networks including junction modeling. *DTA Symposium.* Takayama, Japan.

Rao, L., Owen, L., & Goldsman, D. (1998). Development and Application of a validation framework for traffic simulation models. *Winter Simulation Conference*, (pp. 1079-1086).

Richards P. (1956). Shock waves on the highway. *Operations Research 4*, 42-51.

Salgado, D., Jolovic, D., Martin, P. T., & Aldrete, R. M. (2016). Traffic Microsimulation Models Assessment – A Case Study of International Land Port of Entry. *Procedia Computer Science 83*, 441-448. doi: 10.1016/j.procs.2016.04.207

Seffah, A., Donyaee, M., Kline, R., & Padda, H. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal 14*, 159-178.

Shahpar, A., Aashtiani, H., & Babazadeh, A. (2008). Dynamic penalty function method for the side constrained traffic assignment problem. *Appl. Math. Comput.*, 332-345.

Smith, M. (1993). A new dynamic traffic model and the existence and calculation of dynamic user equilibria on congested capacity-constrained road network. *Transp. Res. Part B Methodol. 27B*, 49-63.

Smits, E.-S., Bliemer, M. C., Pel, A. J., & Arem, B. v. (2015). A family of macroscopic node models. *Transportation Research Part B: Methodological, Volume 74*, 20-39. doi:https://doi.org/10.1016/j.trb.2015.01.002

Taale, H. (2008). *Integrated Anticipatory Control of Road Networks.* Delft, The Netherlands: TRAIL Thesis Series nr. T2008/15.

Taale, H., Westerman, M., Stoelhorst, H., & van Amelsfort, D. (2004). Regional and sustainable traffic management in The Netherlands: methodology and applications. *European Transport Conference* (pp. 1-15). Strasbourg, France: Association for European Transport.

Tampère, C. M., Corthout, R., Cattrysse, D., & Immers, L. H. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B, 45*, 289-309.

TRB, T. R. (2000). *Highway Capacity Manual.* Washington D.C., U.S.A: National Research Council.

Troutbeck, R., & Brilon, W. (2002). Unsignalized Intersection Theory. In *Traffic Flow Theory, A State of the Art Report.* Retrieved January 17, 2005, from www.tfhrc.gov/its/tft/tft.htm

US Department of Transportation. (2004). *Traffic Analysis Toolbox Volume II: Decision Support Methodology for selecting Traffic Analysis Tools.* McLean, Virginia: Federal Highway Administration, Research, Development and Technology, Turner-Fairbank Highway Research Center,.

Van Aerde, M. (1995). Single Regime Speed-Flow-Density Relationship for Congested and Uncongested Highways. *74th TRB Annual Conference, Paper No. 95080.*

van der Gun, J., Pel, A., & van Arem, B. (2017). Extending the Link Transmission Model with non-triangular fundamental diagrams and capacity drops. *Transp. Res. Part B Methodol. 98*, 154-178.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers, Part II 1*, 352-362.

Whitham, G. (1974). *Linear and nonlinear waves.* Wiley, New York.

Wikipedia. (2021, September 24). *Emma (given name)*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Emma_(given_name)

Witteveen+Bos, & Taale, H. (2020, March 16). MARPLE - Beschrijving en handleiding Version - 19e. Rijkswaterstaat.

Yperman, I. (2007). *The Link Transmission Model for Dynamic Network Loading.* Leuven, Belgium: PhD Thesis, Katholieke Universiteit Leuven.

Ypperman, I., Logghe, S., & Immers, B. (2005). The link transmission model: an efficient implementation of the kinematic wave theory in traffic networks. *Advanced OR and AI methods in Transportation*, 122-127.

Zhang, H. (2002). A non-equilibrium traffic model devoid of gas-like behavior. *Transporation Research Part B 36*, 275-290.

Appendix-A: Results of theoretical testing

1. Flow propagation model

The first part of testing involves the results from link propagation, under various scenarios as described below. The test networks mainly adopted from (FakhraeiRoudsari, Huang, & Tampère, 2015) and (Raadsen, Mein, Schilpzand, & Brandt, 2010).

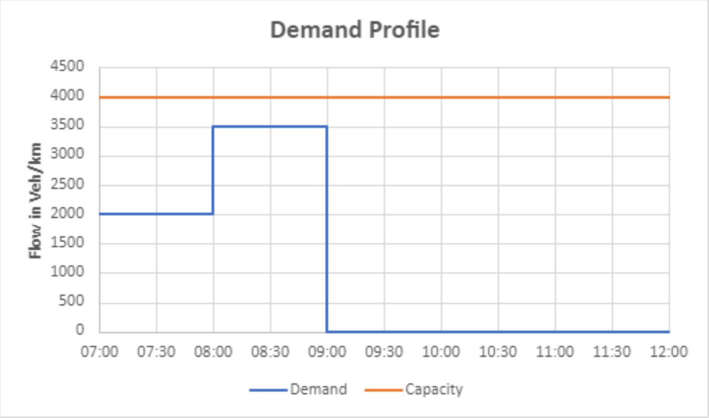
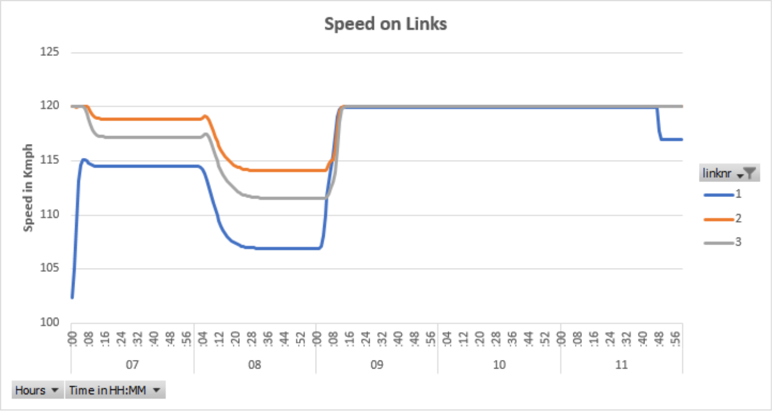
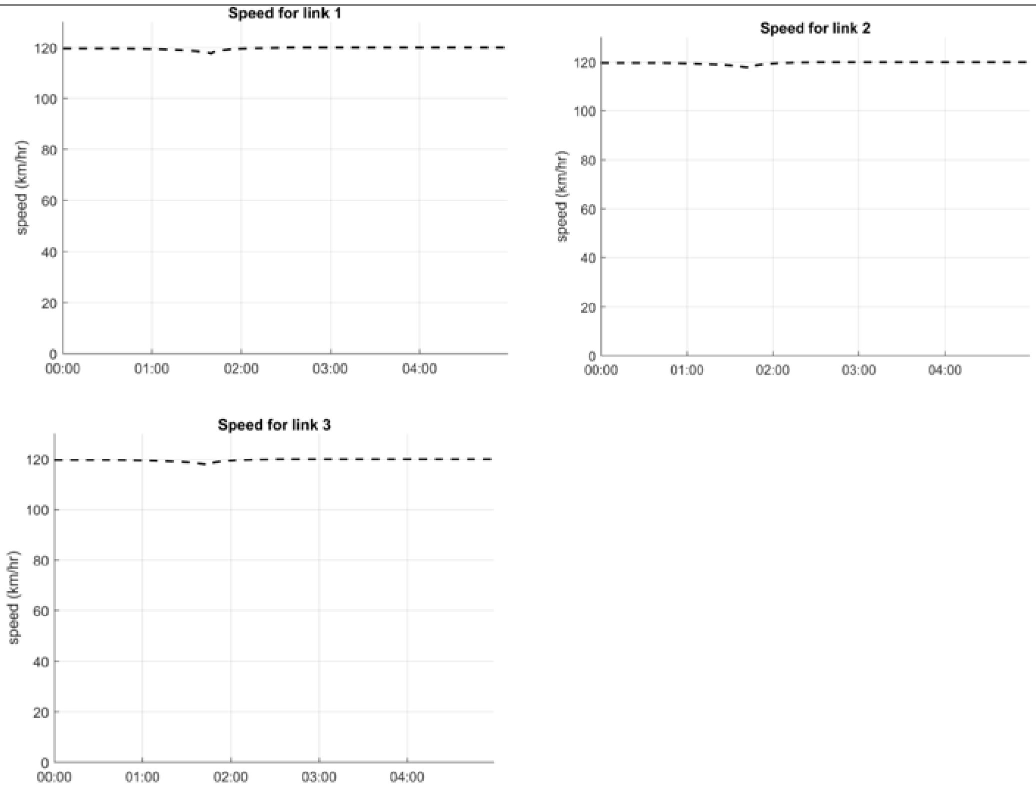
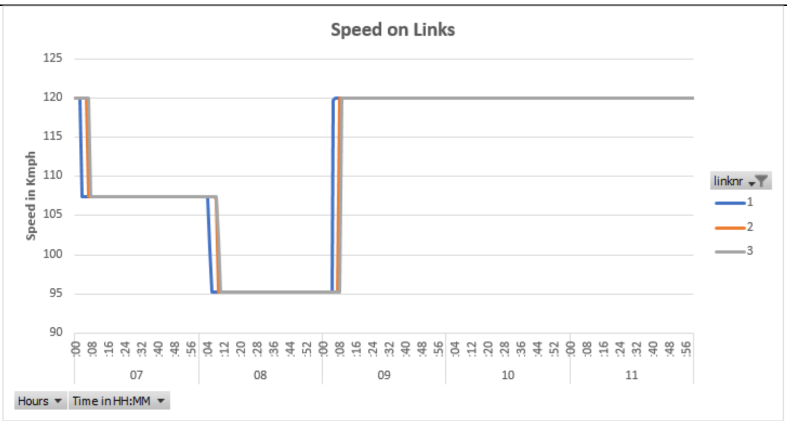
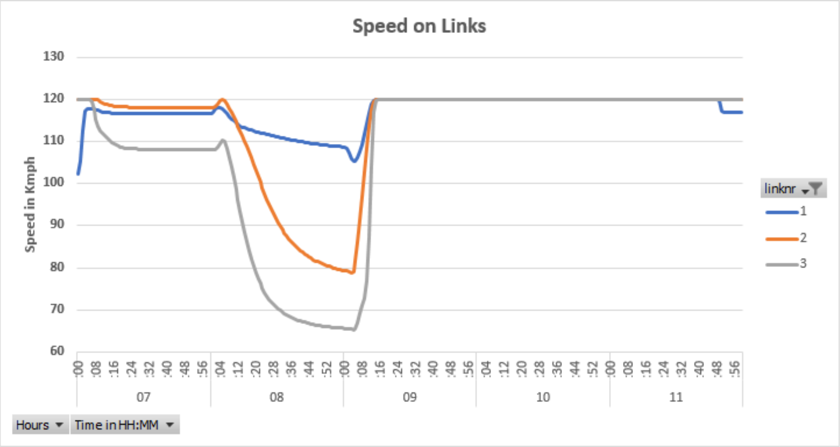
<table>
<tr><th>Test ID</th><th>1.1.1</th></tr>
<tr><td>Test Network</td><td>
</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to check the uninterrupted propagation behavior of the links. Throughout the simulation the all the links are in undersaturated conditions.</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td>Propagation - Link Flows</td></tr>
<tr><td>Supply Properties</td><td>Table 8: Network properties for Test No. 1.1.1
<table>
<tr><th></th><th>Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3</td><td>4000</td><td>7</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Connector_1</td><td>6</td><td>4600</td><td>5</td><td>90</td><td>50</td><td>2</td></tr>
<tr><td>Connector_2</td><td>7</td><td>4600</td><td>5</td><td>90</td><td>50</td><td>2</td></tr>
</table></td></tr>
<tr><td>Demand Properties</td><td>

Figure 30: Demand profile for Test No 1.1.1
5-hour simulation of demand. 2 hours of network loading and 3 hours of emptying the network.</td></tr>
<tr><td>Expectation</td><td>Uninterrupted flow is expected here in all links with speed drop expected from 08:00 to 09:00 due to increase in link saturation value (demand<Capacity, at all times).</td></tr>
</table>

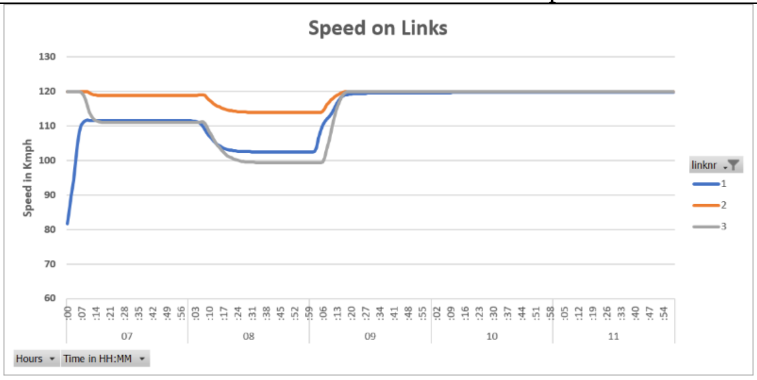
Results - StreamLine: MaDAM	 *Figure 31: Speed values on the corridor links_1.1.1_StreamLine: MaDAM* The results are in line with the expectation. A speed drop can be observed in all links during higher demand as a consequence of saturation. Link 1 and Link 3 have lower speeds compared to Link 2, as they are adjacent to the connector links having a lower free-flow speed. The second order effects are modelled accurately here. However, the drop in speeds, especially in link 2 is close to free speed, which is quite unrealistic.
Score in EMMa	**3**
Results - MARPLE	 *Figure 32:Speed values on the corridor links_1.1.1_MARPLE* As the link propagation of MARPLE is based on travel time functions and do not account of vehicle hysteresis as is the case with a second-order CTM in MaDAM, we will not expect the link speeds varying according to the connector speeds. However, a minor drop in speeds is observed

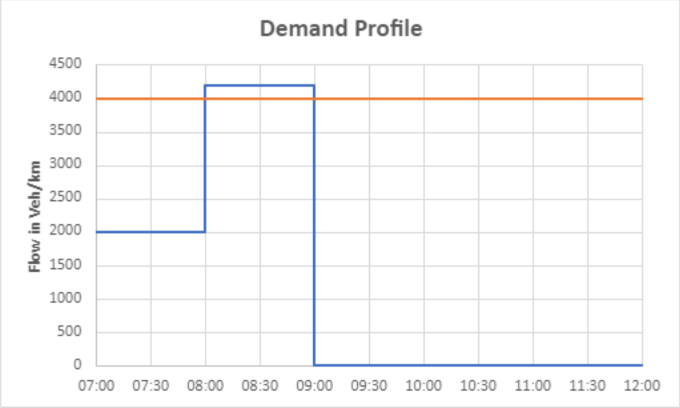
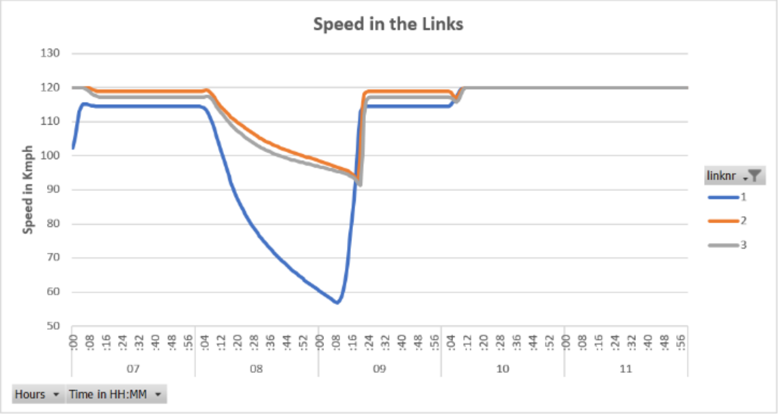
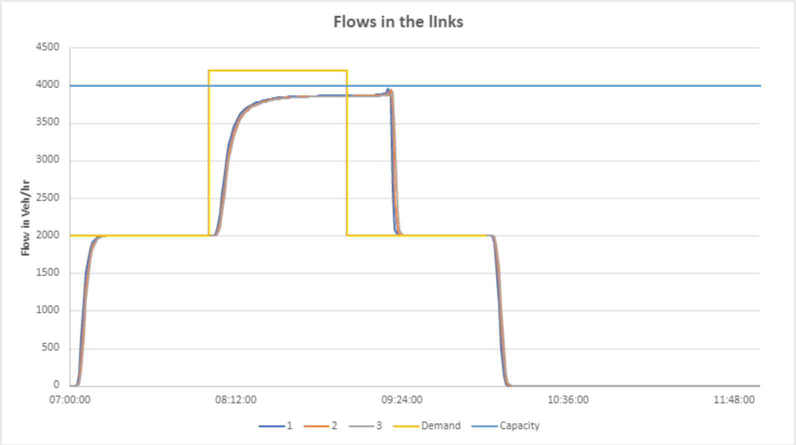
	as a direct consequence of the increase in flow for all three links. However, the drop in speed is very less, and for an instance of time which is unrealistic, as we cannot expect the vehicles to move at free speed (or near free speed ~ 115 Kmph), even when the saturation is 88%.
Score in EMMa	**2**
Results-StreamLine: eGLTM	 *Figure 33: Speed values on the corridor links_1.1.1_StreamLine: eGLTM* The lack of second order effects is evident from the results. The speed drops are sudden and the lower connector speed does not influence the speeds in the main links. The speeds drop down in 08-09 interval as consequence of saturation. However, unlike both the other models, the speed drops are more severe for all the links, which can be expected due to high link saturation value.
Score in EMMa	**4**
Test ID	**1.1.2**
Test Network	Same as 1.1.1
Test Description	Similar to the previous test, the objective of this test is to check the uninterrupted propagation behavior of the links. Difference from the previous test is a variation in the Anticipation parameter- “Nue” to a very high value of 140, as opposed to default value of 35. Throughout the simulation, all the links are in undersaturated conditions. The test is specific to second-order models, to understand the sensitivity of the additional speed parameters.
MoP Evaluated -Link to EMMa	*Propagation - Link Flows*
Supply Properties	Same as 1.1.1, variation only in the anticipation parameter- “Nue” = 140
Demand Properties	Same as 1.1.1
Expectation	Uninterrupted flow is expected here in all links with speed drop expected from 08:00 to 09:00 due to increase in link saturation value (demand<Capacity, at all times). A higher drop in speed values for the links adjacent to the connector links are expected, due to larger Nue value. Furthermore, the larger effects of anticipation from links downstream are expected compared to the convection from upstream links.

<table>
<tr><td>Results - StreamLine: MaDAM</td><td>

<i>Figure 34: Speed values on the corridor links_1.1.2_StreamLine: MaDAM</i>

An increase of value of anticipation parameter, has resulted in smaller difference links 1 and 2, in comparison to links 1 and 3 (stronger influence of downstream anticipation compared to upstream connector link). In line with the same logic, link 3 is influenced greatly now by the connector link upstream to the destination with lower speed, which can be visualized in Figure 34. The high value of vehicle speeds during a denser saturation in Link 2 still exists in this test result.</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
<tr><td>Results - MARPLE</td><td><i>Test cannot be performed as model is not second-order based.</i></td></tr>
<tr><td>Score in EMMa</td><td>NA</td></tr>
<tr><td>Results- StreamLine: eGLTM</td><td><i>Test cannot be performed as model is not second-order based.</i></td></tr>
<tr><td>Score in EMMa</td><td>NA</td></tr>
<tr><td>Test ID</td><td>1.1.3</td></tr>
<tr><td>Test Network</td><td>Same as 1.1.1</td></tr>
<tr><td>Test Description</td><td>Similar to the previous test, the objective of this test is to check the uninterrupted propagation behavior of the links. Difference from the previous test is a variation in the free-flow speed of the connector link with default values of speed terms. Throughout the simulation, all the links are in undersaturated conditions. The test is specific to second-order models, to understand the sensitivity of link speeds to connecting links upstream and downstream.</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td><i>Propagation - Link Flows</i></td></tr>
<tr><td>Supply Properties</td><td><i>Table 9: Network properties for Test No. 1.1.3</i>
<table>
<tr><th></th><th>Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3</td><td>4000</td><td>7</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Connector_1</td><td>6</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
<tr><td>Connector_2</td><td>7</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
</table></td></tr>
</table>

Demand Properties	Same as 1.1.1
Expectation	Uninterrupted flow is expected here in all links with speed drop expected from 08:00 to 09:00 due to increase in link saturation value (demand<Capacity, at all times). A larger drop in speed is expected in links 1 and 3 as the connector links have lower speeds.
Results - StreamLine: MaDAM	 *Figure 35: Speed values on the corridor links_1.1.3_StreamLine: MaDAM* Larger difference is witnessed between link speeds in Link 1 and Link 3 vs Link 2. However, link speeds in 2 reaches free flow speeds at higher demand which is still unrealistic. When compared to test no 1.1.1, link speeds of 3 seems lower than that of 1. This could be the influence of the additional speed term or specifically the anticipation term in combination with the convection term which has dragged down the vehicle speeds.
Score in EMMa	**3**
Results - MARPLE	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Results-StreamLine: eGLTM	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Test ID	**1.2.1**
Test Network	Same as 1.1.1
Test Description	The objective of the test is to check the demand propagation when the link flows are oversaturated at a specific time period in simulation.
MoP Evaluated -Link to EMMa	*Propagation - Queuing*
Supply Properties	Same as 1.1.1

Demand Properties	 *Figure 36: Demand profile for Test No 1.2.1* 5-hour simulation of demand. 2 hours of network loading and 3 hours of emptying the network
Expectation	The expectation would be that the oversaturation in the links would cause queue in the origin (non-connector link closest to the origin).
Results - StreamLine: MaDAM	 *Figure 37: Speed values on the corridor links_1.2.1_StreamLine: MaDAM*  *Figure 38: Flow values on the corridor links_1.2.1_StreamLine: MaDAM* The speed drop in Link 1 during the period of oversaturation is evident from the speed values shown in Figure 37. Furthermore, queuing in the link 1 causes the outflow of all links to the extent of capacity. We can observe from Figure 38, that the outflow values are below the capacity. The speed dynamics in the second order CTM, incorporates this phenomenon of capacity drop.

<table>
<tr>
<td></td>
<td>However, model does not offer any quantitative means to control this drop in capacity for the outflow of downstream links from a bottleneck.</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>4</td>
</tr>
<tr>
<td>Results - MARPLE</td>
<td>

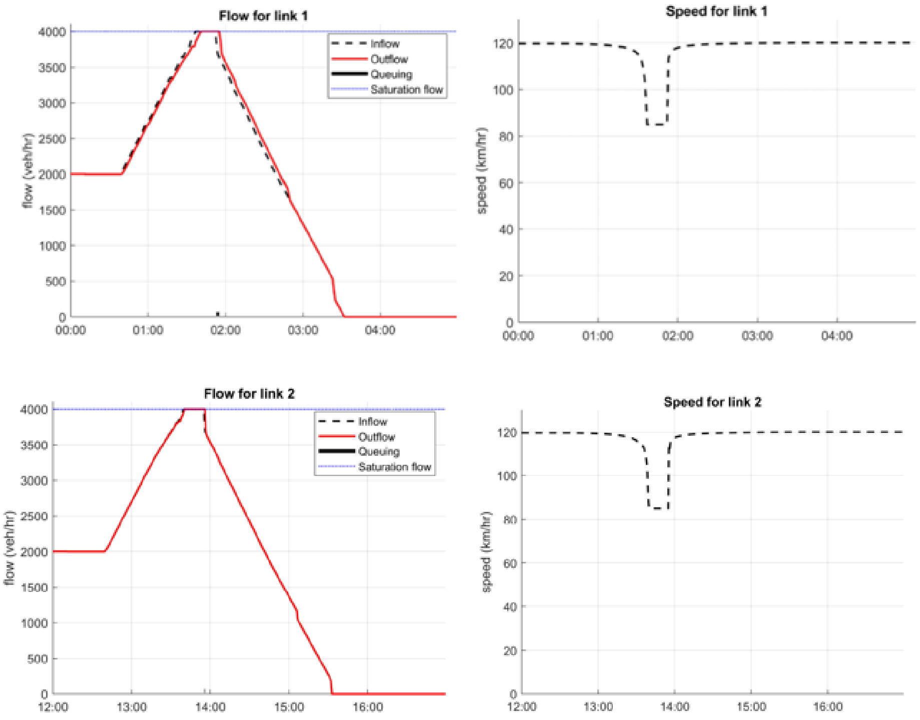


Figure 39: Speed and flow values on the corridor links_1.2.1_MARPLE

A drop in speed is observed in some time instances during the second hour of the simulation, due to congestion in the at the origin (demand>saturation flow). The drop is closer to 85km/hr as can be observed speed plots as the congested speeds are set to 85Km/hr in MARPLE for an uncontrolled link. The speed drop is also gradual, as can be observed through the plots. A capacity drop as visualized in case of the 2nd Order CTM model cannot be viewed in this case. From the flow plot, it can be observed that the outflow exactly matches the link capacity.

</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>2</td>
</tr>
<tr>
<td>Results- StreamLine: eGLTM</td>
<td>

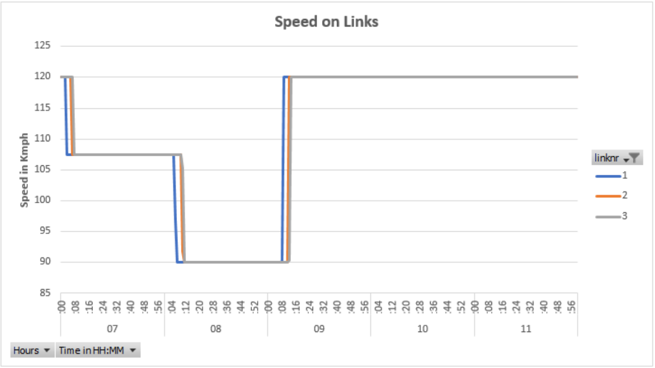


Figure 40: Speed values on the corridor links_1.2.1_StreamLine: eGLTM

</td>
</tr>
</table>

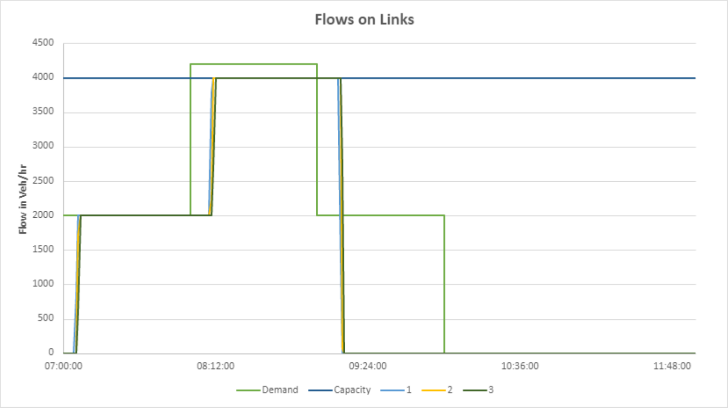
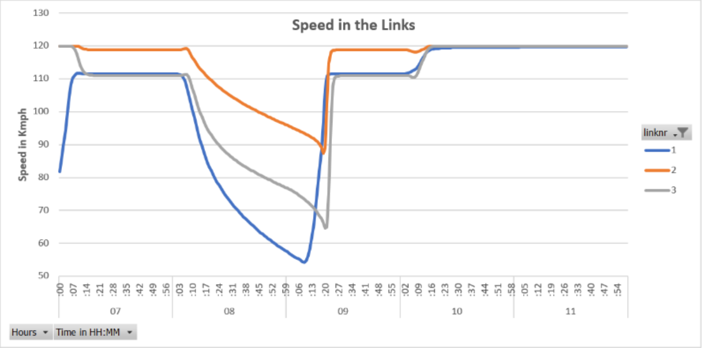
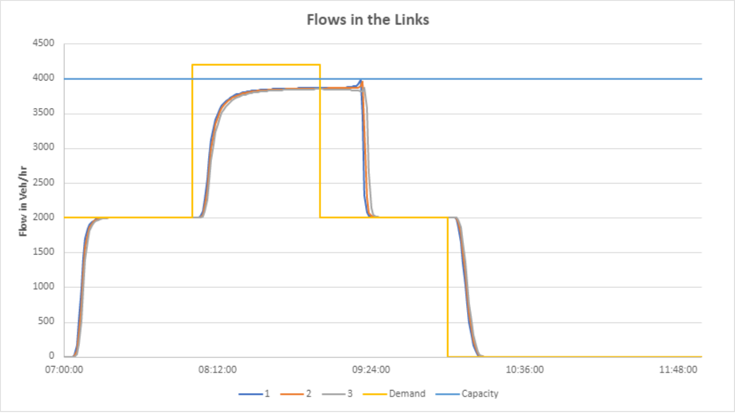
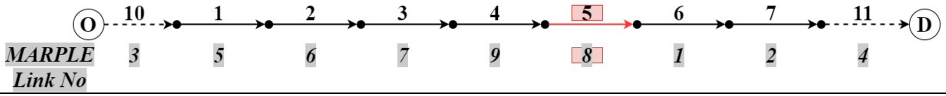
<table>
<tr><td></td><td>

Figure 41: Flow values on the corridor links_1.2.1_StreamLine: eGLTM

Overall, the behavior is as per expectation. However, the extent of decrease in speed is not as per expectation. For an oversaturated condition, a larger speed drop will be expected. The sudden increase and decrease of speeds and flows are also evident from Figure 40 and Figure 41, due to the absence of hysteresis. Capacity drop is not observed from the results.</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
<tr><td>Test ID</td><td>1.2.2</td></tr>
<tr><td>Test Network</td><td>Same as 1.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of the test is to check the demand propagation when the link flows are oversaturated at a specific time period in simulation, when the connector links from the origin have a lower free-flow speed. The second order effect of links during queuing is checked and evaluated through this test.</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td><i>Propagation – Queuing</i></td></tr>
<tr><td>Supply Properties</td><td><i>Table 10: Network properties for Test No. 1.2.2</i>
<table>
<tr><th></th><th>Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3</td><td>4000</td><td>7</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Connector_1</td><td>6</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
<tr><td>Connector_2</td><td>7</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
</table></td></tr>
<tr><td>Demand Properties</td><td>Same as 1.1.3</td></tr>
<tr><td>Expectation</td><td>The expectation would be that the oversaturation in the links would cause queue in the origin (non-connector link closest to the origin), with speeds on Link 2 adjusting to the lower speeds downstream (consequence of the anticipation term).</td></tr>
</table>

Results - StreamLine: MaDAM	 *Figure 42: Speed values on the corridor links_1.2.2_StreamLine: MaDAM*  *Figure 43: Flow values on the corridor links_1.2.2_StreamLine: MaDAM* Difference between the link speeds in those near the connector links are more compared to link no.2 speed, as shown in Figure 42. However, it is not practical for the speeds of the cars to be 60kmph especially during congestion. Therefore, a question on the calibration of the speed terms parameters. Furthermore, looking at the flow values in Figure 43, it can be observed that at the peak period, an effect of capacity drop can be observed, as the flow values post congestion, is less than the capacity. however, the effect cannot be controlled in streamline. As observed in the flow's chart, the congestion is built up at the origin, where the phenomenon of capacity drop is observed.
Score in EMMa	**3**
Results - MARPLE	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Results-StreamLine: eGLTM	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Test ID	**1.3.1**
Test Network	

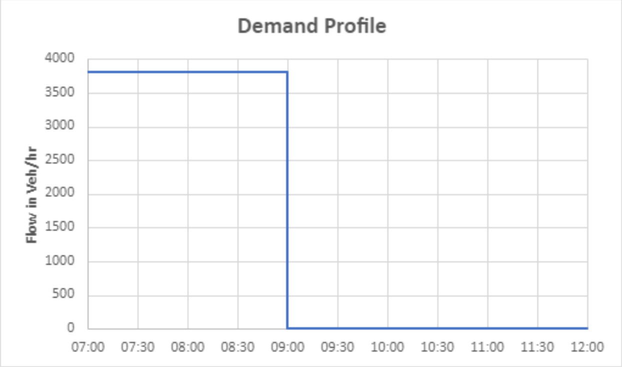
<table>
<tr><td></td><td>In comparison with the previous tests, the network has been dissagregated into more links. The same was done inorder to capture the dynamics upstream and downstream of the bottleneck link nr 5. Please note that for MARPLE, the link nrs are different from StreamLine: MaDAM and eGLTM, as provided in the text in grey background. The reader is refered to this network image while reading the results for MARPLE. All the charts and result discussion related to MARPLE will be based on this link numbering.</td></tr>
<tr><td>Test Description</td><td>The objective of the test is to understand the queuing and propagation behavior during activation of a stationary bottleneck link, under a constant demand propagation.</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td>Propagation – Queuing</td></tr>
<tr><td>Supply Properties</td><td>Table 11: Network properties for Test No. 1.3.1
<table>
<tr><th></th><th>Link Nr</th><th>MARPLE Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3, 4,6,7</td><td>5,6,7,9,1,2</td><td>4000</td><td>1</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Corridor link (bottleneck)</td><td>5</td><td>8</td><td>3000</td><td>1</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Connectors</td><td>10,11</td><td>3,4</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
</table></td></tr>
<tr><td>Demand Properties</td><td>

Figure 44: Demand profile for Test No 1.3.1

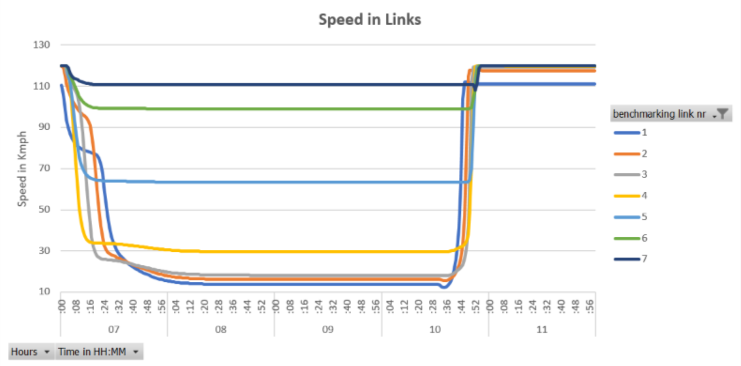
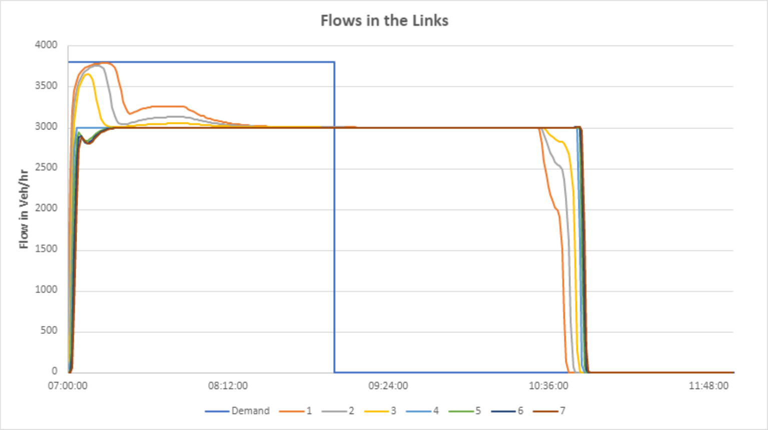
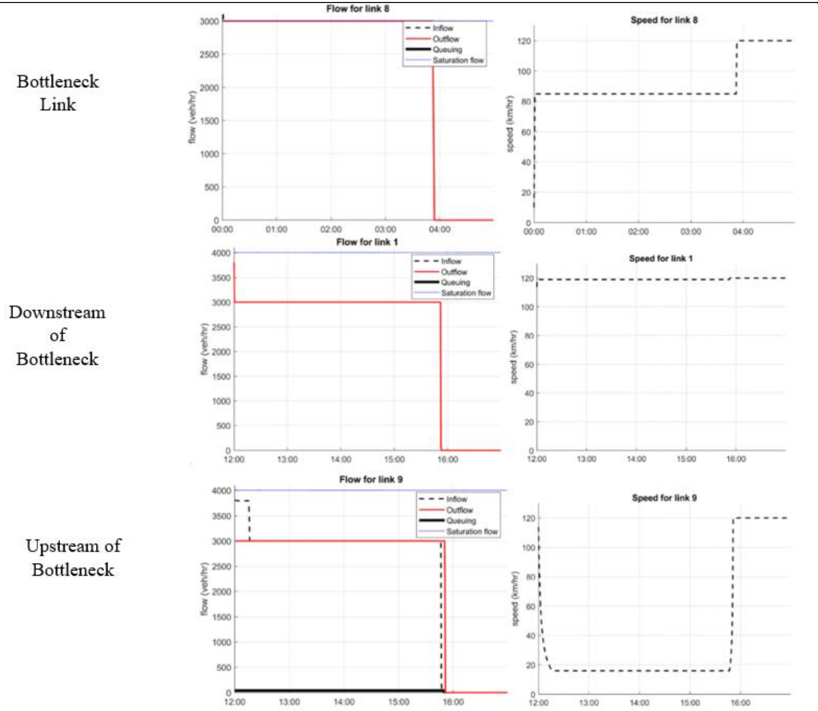
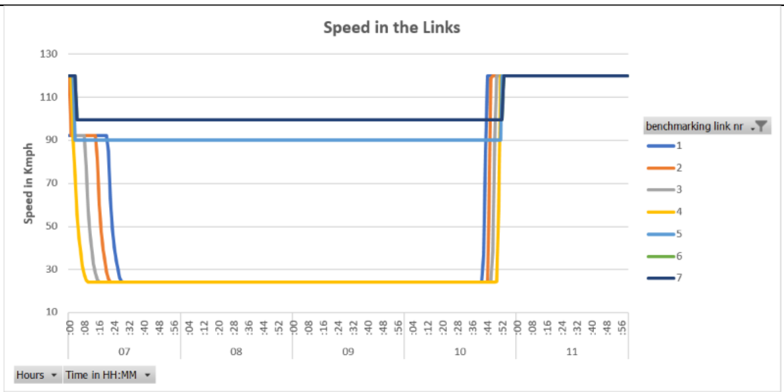
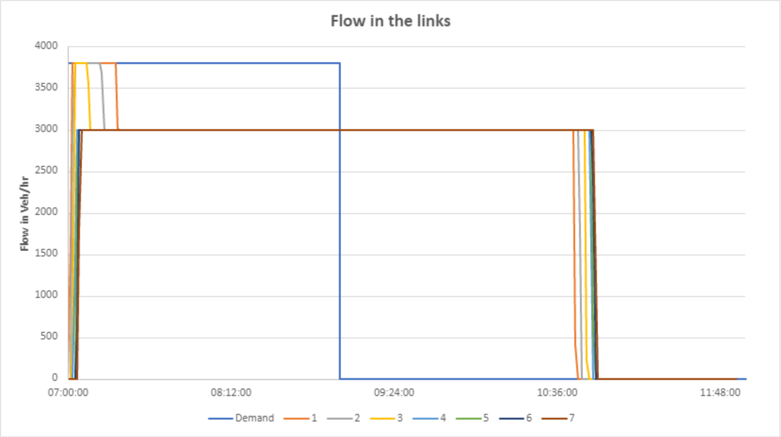
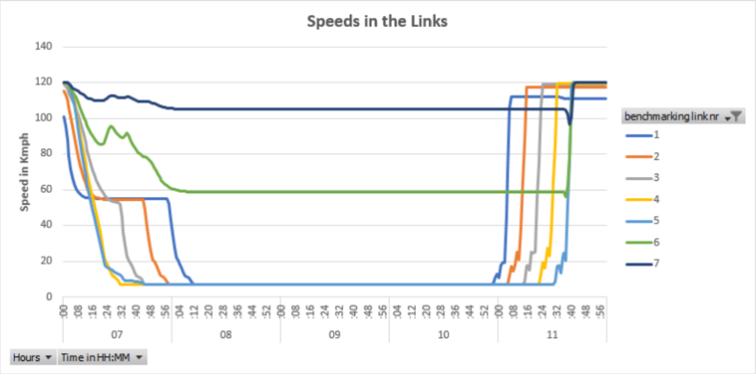
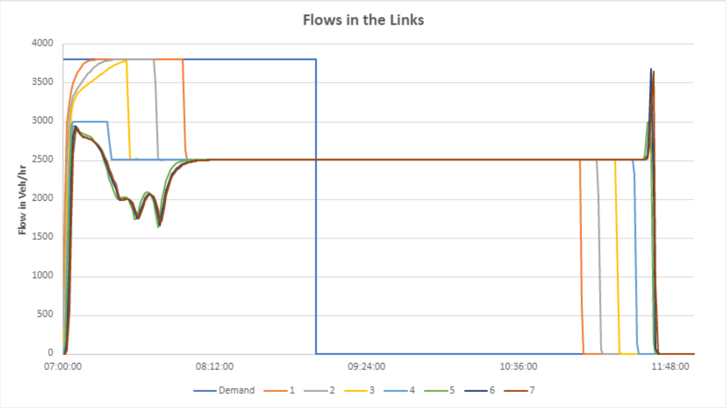
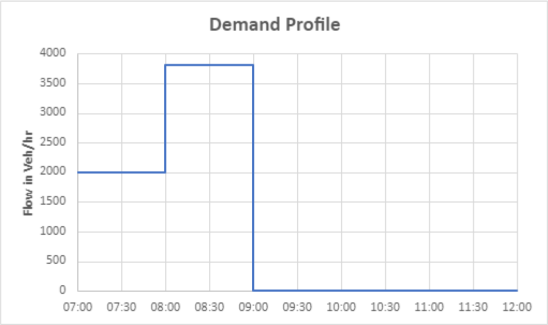
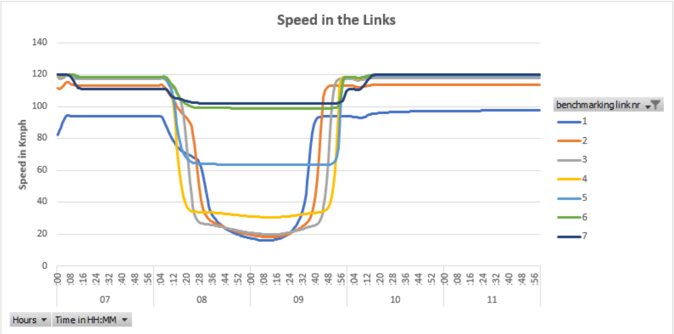
5 hr simulation with 2 hrs of constant network loading of value 3800 veh/hr.</td></tr>
<tr><td>Expectation</td><td>A Bottleneck is expected to form at Link 5 location, as demand (3800 veh/hr) exceeds its capacity (3000 veh/hr). The subsequent queuing in the preceding link 4, is expected to cause a spillback in the links upstream. Downstream of the bottleneck, the link outflow will be expected to match the capacity of the bottleneck link, during the propagation time period.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

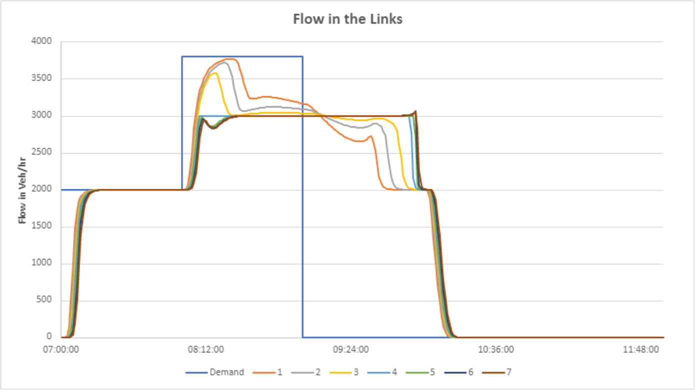
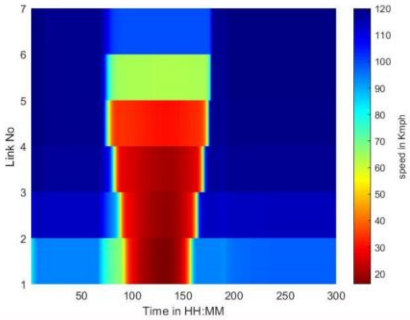
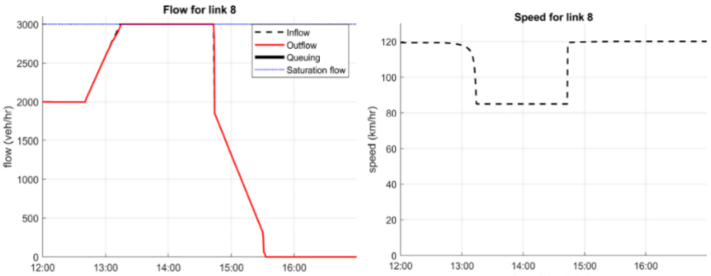
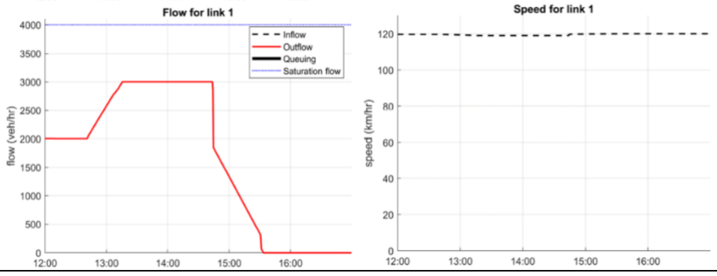
Figure 45: Speed values on the corridor links_1.3.1_StreamLine: MaDAM</td></tr>
</table>

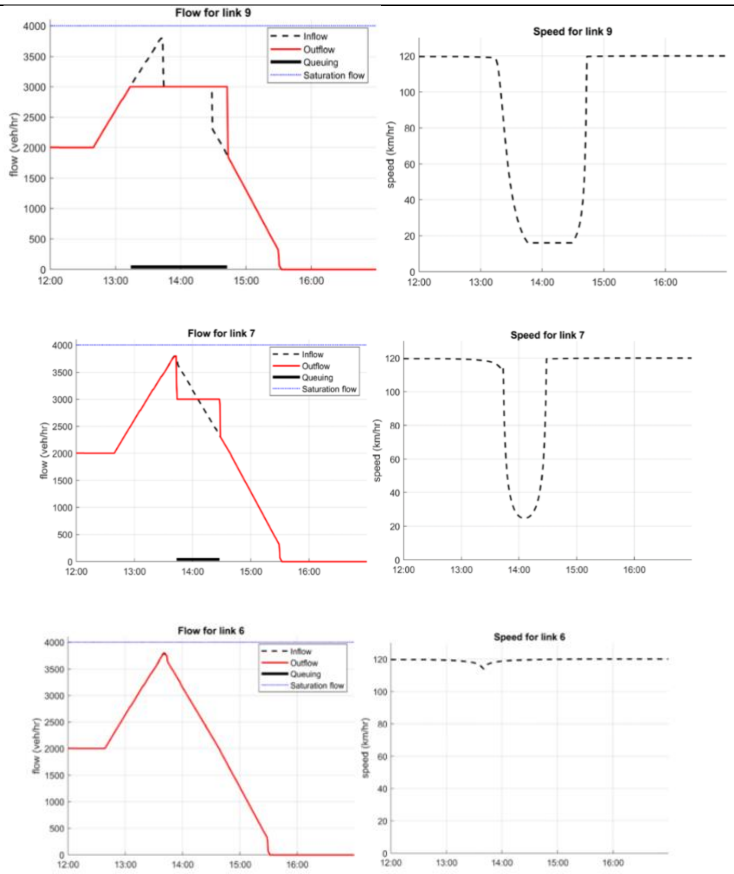
	 *Figure 46: Flow values on the corridor links_1.3.1_StreamLine: MaDAM* Link speeds are adjusted to be the lowest for the first link during the activation of the bottleneck, as witnessed in Figure 45. This can be expected due to the interplay of various factors including the convection and anticipation speed terms, effect of congestion and thereby spillback. Overall, the queuing and the spillback takes place as expected during the duration of the demand loading. However, from Figure 46, the results of the link outflows are contradictory to that observed in the previous test cases for MaDAM. After the resolution of the queue, the capacity drop is not observed for the any of the links. Thus, a reliable model result is not showcased for capacity drop. The roughness factor described in Section 3.2.1 and Formula 10 is calculated using this test case. The roughness factor for link upstream of the bottleneck (Link Nr 4) is selected for this. The values are as provided: • *Distribution of vehicle speeds - Arithmetic Average Height (Ra) = 34* • *Distribution of traffic flows - Arithmetic Average Height (Ra) = 1088*
Score in EMMa	**3**
Results - MARPLE	 *Figure 47: Speed and flow values on the corridor links_1.3.1_MARPLE*

	The modelled results showed an anomalous speed fluctuation initially for a propagation time step of 60s. This was eliminated when the time step size is reduced to 30s as shown in Figure 47. This means the model is giving inconsistent results while adjusting short links (Link travel time with free-flow speed < length of the simulation time step). The bottleneck flow is consistent with expectation and inflow values match the outflow and saturation flow values throughout the period of demand. As a result of blocking back, queuing is observed and is plotted in the graph. The queue dissipation is really slow and propagation is present consistently during the 3rd hour and 4th hour of the simulation. Speed values are consistent with the expectation. Upstream link's flow and speed graph shows consistent spillback behavior from the bottleneck. The dissolution period after the demand drops to 0, takes almost an hour for emptying the network. The roughness factor described in Section 3.2.1 and Formula 10 is calculated using this test case. The roughness factor for link upstream of the bottleneck (Link Nr 9, for MARPLE) is selected for this. The values are as provided: - *Distribution of vehicle speeds - Arithmetic Average Height (Ra) = 45* - *Distribution of traffic flows - Arithmetic Average Height (Ra) = 18*
Score in EMMa	**2**
Results-StreamLine: eGLTM	 *Figure 48: Speed values on the corridor links_1.3.1_StreamLine: eGLTM*  *Figure 49: Flow values on the corridor links_1.3.1_StreamLine: eGLTM* The speed drop is maximum for the links upstream from the bottleneck, illustrated from the speed graph in Figure 48. The phenomenon of capacity drop or a smooth variation in speeds and flows are not observed, which can be understood because of the absence of 2nd order effects. The link outflow values in Figure 49 show spillback into the links

	upstream from the congestion. The roughness factor described in Section 3.2.1 and Formula 10 is calculated using this test case. The roughness factor for link upstream of the bottleneck (Link Nr 4) is selected for this. The values are as provided: - *Distribution of vehicle speeds - Arithmetic Average Height (Ra) = 35* - *Distribution of traffic flows - Arithmetic Average Height (Ra) = 1089*
Score in EMMa	**3**
Test ID	**1.3.2**
Test Network	Same as 1.3.1
Test Description	To test the propagation behavior in the event of a stationary bottleneck under constant demand. The difference between the previous test case is that the anticipation parameter nue is reduced to null, to observe the influence of anticipation term in link propagation. The test is restricted to second order models.
MoP Evaluated -Link to EMMa	*Propagation – Queuing*
Supply Properties	Same as 1.3.1, except the anticipation parameter nue = 0.
Demand Properties	Same as 1.3.1
Expectation	We would expect traffic to propagate until it is inside the bottleneck link 5 (since it didn't anticipate to it) and then the relaxation term would push speeds, and herewith flow, downstream. Severe congestion would form inside the bottleneck with density diverging to infinity, as upstream traffic does not anticipate this queue and keeps on flowing in.
Results - StreamLine: MaDAM	 *Figure 50: Speed values on the corridor links_1.3.2_StreamLine: MaDAM*  *Figure 51: Flow values on the corridor links_1.3.2_StreamLine: MaDAM* However, looking at the flows and speeds in Figure 50 and Figure 51, the software overrules the misspecification by the user for the wrong nue value, by avoiding infinite density values. The speed drop and the flow drop in link 5, is a direct consequence of bottleneck formation. This

	shows a reliable modeling of the upstream congestion propagation by Streamline in case of a stationary bottleneck.
Score in EMMa	**3**
Results - MARPLE	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Results-StreamLine: eGLTM	*Test cannot be performed as model is not second-order based.*
Score in EMMa	NA
Test ID	**1.4.1**
Test Network	Same as 1.3.1
Test Description	The objective of the test is to understand the queuing and propagation behavior during activation of a stationary bottleneck link, under a peak demand (Demand > Bottleneck capacity) for a specific period within the demand simulation.
MoP Evaluated -Link to EMMa	*Propagation – Queuing*
Supply Properties	Same as 1.3.1
Demand Properties	 *Figure 52: Demand profile for Test No 1.4.1*
Expectation	Queuing in Link 4 is expected in during the peak demand. The subsequent queue is expected to spillback into the links upstream.
Results - StreamLine: MaDAM	 *Figure 53: Speed values on the corridor links_1.4.1_StreamLine: MaDAM*

	 *Figure 54: Flow values on the corridor links_1.4.1_StreamLine: MaDAM*  *Figure 55: Space-time diagram of the corridor links_ 1.4.1_StreamLine: MaDAM* Capacity drop is modelled reliably as shown in Figure 54. The results are in line with the expectation along with the propagation characteristics on each link. The same case is represented through the space-time plot which shows that the congestion propagation begins at peak demand and spills all the way to the origin. The same is resolved when the demand value drops below the capacity and the bottleneck is resolved from the tail of the queue.
Score in EMMa	**4**
Results - MARPLE	Bottleneck Link  Downstream of Bottleneck 

	Upstream of Bottleneck  *Figure 56: Flow and Speed values on the corridor links_1.4.1_MARPLE* The outflow max value for the bottleneck link is in line with the expectation. The speed flow values are consistent with the expectation, with the speed in congestion dropping to 85 km/hr during the peak hour demand as shown in Figure 56. Outflow of the links downstream of the bottleneck is also restricted to the capacity of the bottleneck link as per expectation. Upstream link speed and flow trends are also consistent with the expectation. It can be observed that the spillback effect is present and the outflow value of the link upstream matches the inflow values of the link under study (Outflow of link 7 same as inflow of link 9). The outflow remains at 3000 veh/hr for the upstream links even though demand values are high as a result of spillback. As expected the queue formation begins on the link directly upstream of the bottleneck. The anomaly related with shortlinks as described in Test 1.3.1, was also encountered in this test and the time step of propagation was accordingly adjusted.
Score in EMMa	**3**

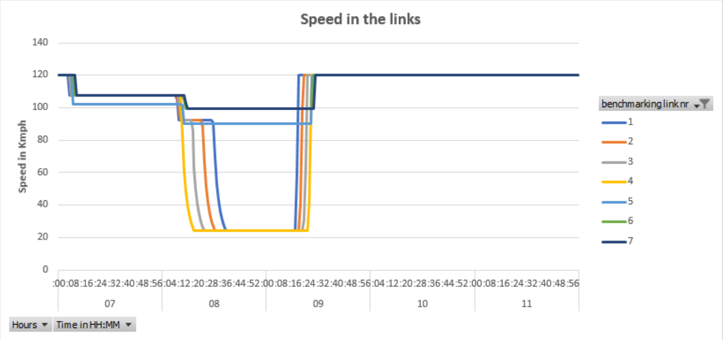
<table>
<tr><td>Results-StreamLine: eGLTM</td><td>


Figure 57: Speed values on the corridor links_1.4.1_StreamLine: eGLTM

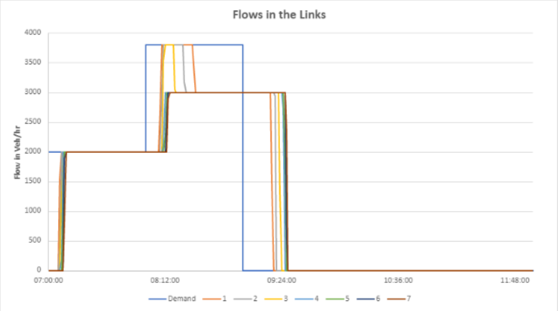
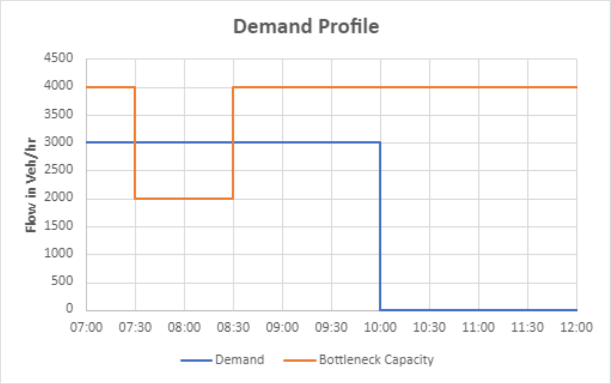
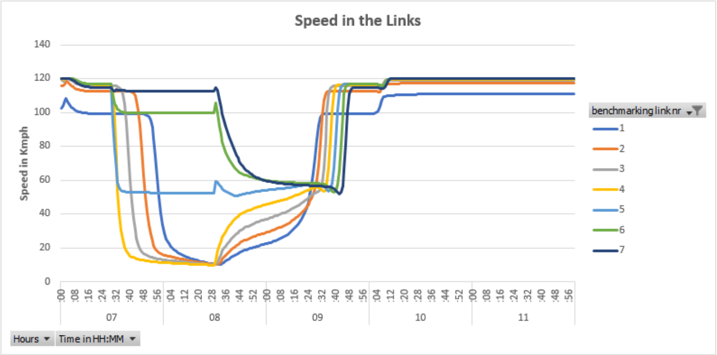
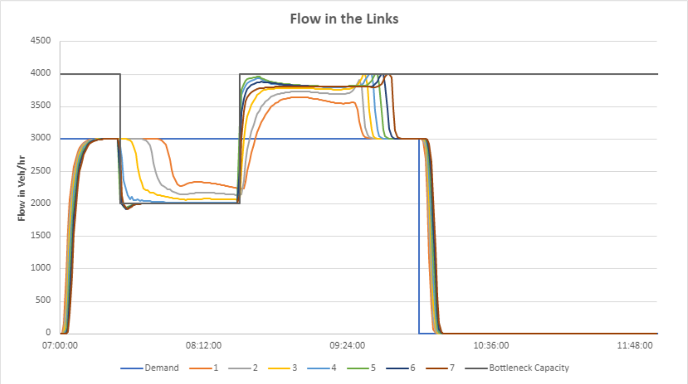
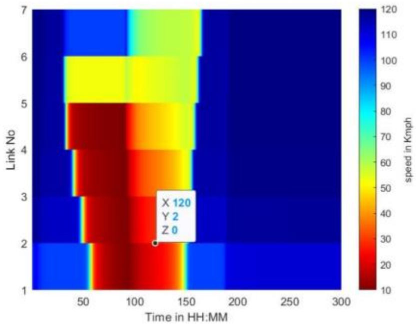


Figure 58: Flow values on the corridor links_1.4.1_StreamLine: eGLTM

The results are in line with the expectation along with he propagation characteristics on each link. The queue formation begins in Link 4, as shown in Figure 58. The same is resolved when the demand value drops below the capacity and the congestion is resolved from the tail of the queue. The recovery occurs rather quickly compared with MaDAM results, mostly due to the absence of hysteresis term.
</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>**Test ID**</td><td>**1.5.1**</td></tr>
<tr><td>Test Network</td><td>Same as 1.3.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to check the influence of a temporary bottleneck in the corridor network. The same is checked by means of an external control for variable capacity of the bottleneck link. The control in the model is used to mimic the following real-world incidents:
• Roadworks (Capacity reduction)
• Variable Speed Limits(VSL) - Speed adjustment
• Weather Changes (Capacity reduction, speed at capacity reduction)</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td>*Propagation - Effect of link-level traffic controls*</td></tr>
<tr><td>Supply Properties</td><td>
Table 12: Network properties for Test No. 1.5.1
<table>
<tr><th></th><th>Link Nr</th><th>MARPLE Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3, 4,6,7</td><td>5,6,7,9,1,2</td><td>4000</td><td>1</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Corridor link (bottleneck)</td><td>5</td><td>8</td><td>3000, 2000 (07:30 - 08:30)</td><td>1</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Connectors</td><td>10,11</td><td>3,4</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
</table>

Link 5 capacity, changed from 3000 to 2000 veh/h for one hour.
</td></tr>
</table>

Demand Properties	 *Figure 59: Demand profile for Test No 1.5.1*
Expectation	The temporary bottleneck is expected to cause queuing in Link 4, and further spillback upstream, during the activation of the bottleneck.
Results - StreamLine: MaDAM	 *Figure 60: Speed values on the corridor links_1.5.1_StreamLine: MaDAM*  *Figure 61: Flow values on the corridor links_1.5.1_StreamLine: MaDAM*  *Figure 62: Space-time diagram of the corridor links_ 1.5.1_StreamLine: MaDAM*

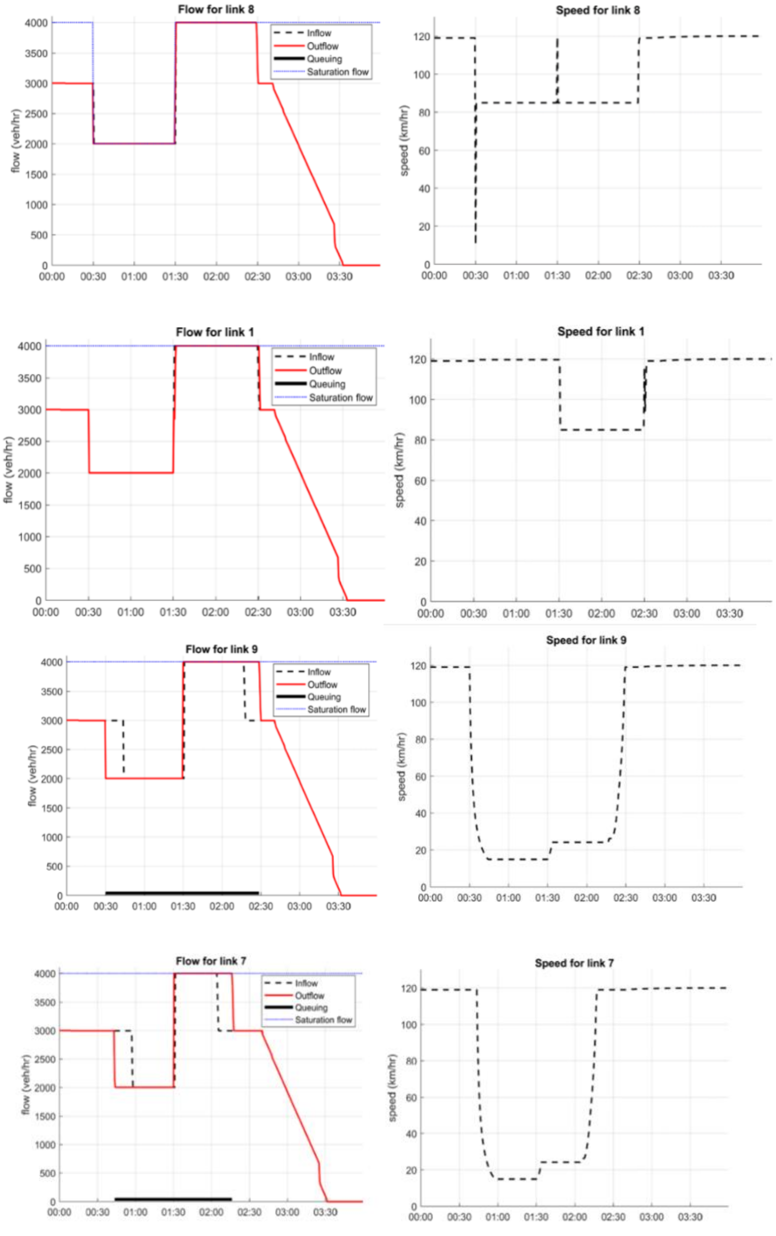
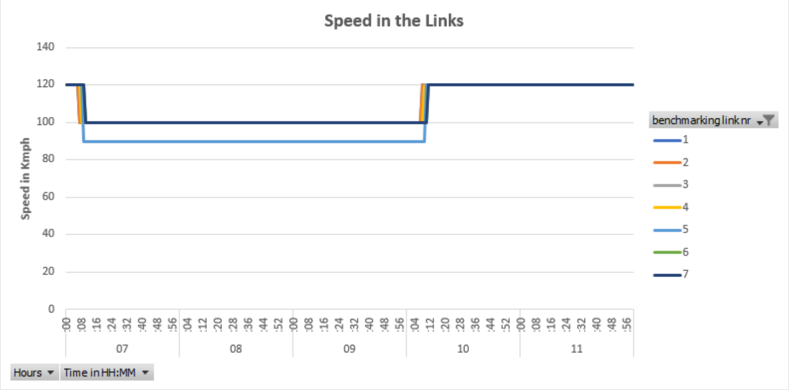
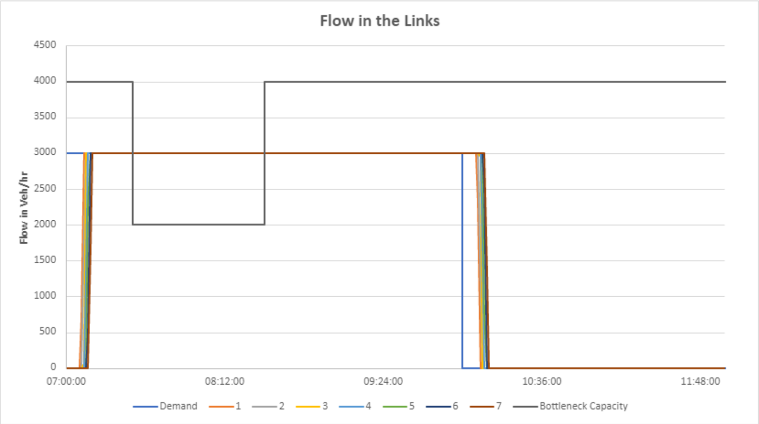
	As expected, the temporary blockade causes the flows and thereby the speeds to drop as an effect of congestion as shown in Figure 60 and Figure 61. When the capacity is restored, we again observe capacity drop phenomenon, consistent with the previous results and our theoretical expectation. As observed from the space-time plot the queue resolution is observed from the front as the capacity is restored beyond 8.30 AM. the speed terms parameters ensure a capacity drop phenomenon at the queue discharge beyond the temporary blockade. However, the capacity drop is not present when the blockade is active, which makes the results uncontrollable.
Score in EMMa	**3**
Results - MARPLE	Bottleneck Link Downstream of Bottleneck Upstream of Bottleneck  *Figure 63: Flow and Speed values on the corridor links_1.5.1_MARPLE* Results are in line with the expectation. The activation of control has reduced the saturation flow from 07:30 to 08:30. The resulting bottleneck has caused blocking back in the links upstream thereby restricting the outflows in those links. Interestingly in the bottleneck link, after the control is deactivated beyond 08:30, a sudden peak is observed in the speed chart, which falls back into a speed drop. This sudden peak may also be an anomaly due to short link(which was indeed the case when tested with smaller propagation time step – results not shown here). The vehicle speeds of the bottleneck and downstream links, when at full capacity (beyond 08:30) as shown in the

	Figure 63, are in free flow state. In that case the speed-flow curve (flow-travel time curve) is at its turning and the speed (speed at capacity) is lower than the maximum speed.
Score in EMMa	**4**
Results-StreamLine: eGLTM	 *Figure 64: Speed values on the corridor links_1.5.1_StreamLine: eGLTM*  *Figure 65: Flow values on the corridor links_1.5.1_StreamLine: MaDAM* Results are not as per expectation. It can be noticed that eGLTM is not sensitive to variable capacity and the outflow does not alter with the change in input capacity. This shows a stark contrast to the results in MaDAM. From discussion with the model developers, it was understood that traffic controls are still to be modelled in the final version. The version of the model tested for this evaluation does not incorporate link controls. The prototype build phase is currently in progress.
Score in EMMa	**0**
Test ID	**1.6.1**
Test Network	Same as 1.3.1
Test Description	The objective of the test is to check the influence of Variable Speed Limit in the corridor network. The same is checked by means of an external control for VSL in the bottleneck link nr 5.
MoP Evaluated -Link to EMMa	*Propagation - Effect of link-level traffic controls*

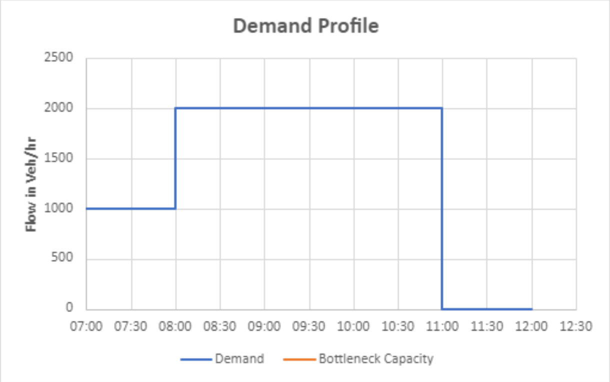
<table>
<tr>
<td>Supply Properties</td>
<td>
Table 13: Network properties for Test No. 1.6.1
<table>
<tr><th></th><th>Link Nr</th><th>MARPLE Link Nr</th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>1,2,3, 4,6,7</td><td>5,6,7,9,1,2</td><td>4000</td><td>1</td><td>120</td><td>90</td><td>2</td></tr>
<tr><td>Corridor link (bottleneck)</td><td>5</td><td>8</td><td>3000,</td><td>1</td><td>120, 80(07:30 - 08:30)</td><td>50, 60(07:30 - 08:30)</td><td>2</td></tr>
<tr><td>Connectors</td><td>10,11</td><td>3,4</td><td>4600</td><td>5</td><td>50</td><td>35</td><td>2</td></tr>
</table>
Link 5 Free flow speed, reduced from 120 to 80 Kmph for one hour.
</td>
</tr>
<tr>
<td>Demand Properties</td>
<td>


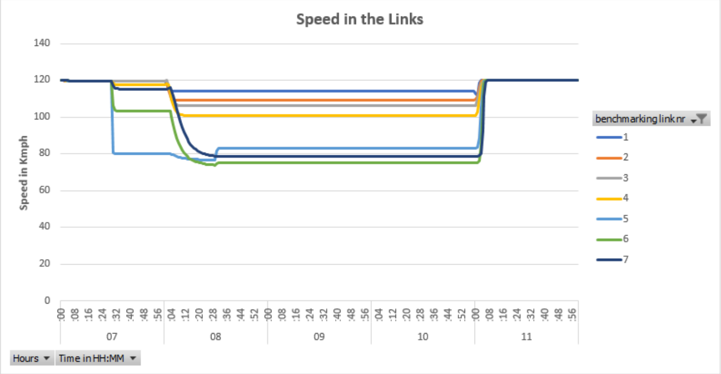
Figure 66: Demand profile for Test No 1.6.1
</td>
</tr>
<tr>
<td>Expectation</td>
<td>The speed values in the bottleneck link are expected to drop during the activation of the bottleneck between 07:30 – 08:30. A consequent drop in flows is also expected during this.</td>
</tr>
<tr>
<td>Results - StreamLine: MaDAM</td>
<td>


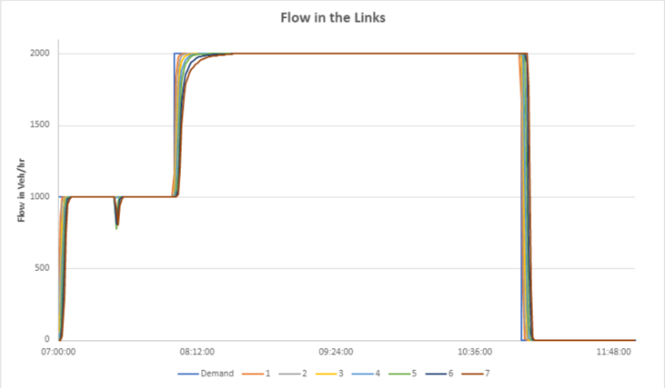
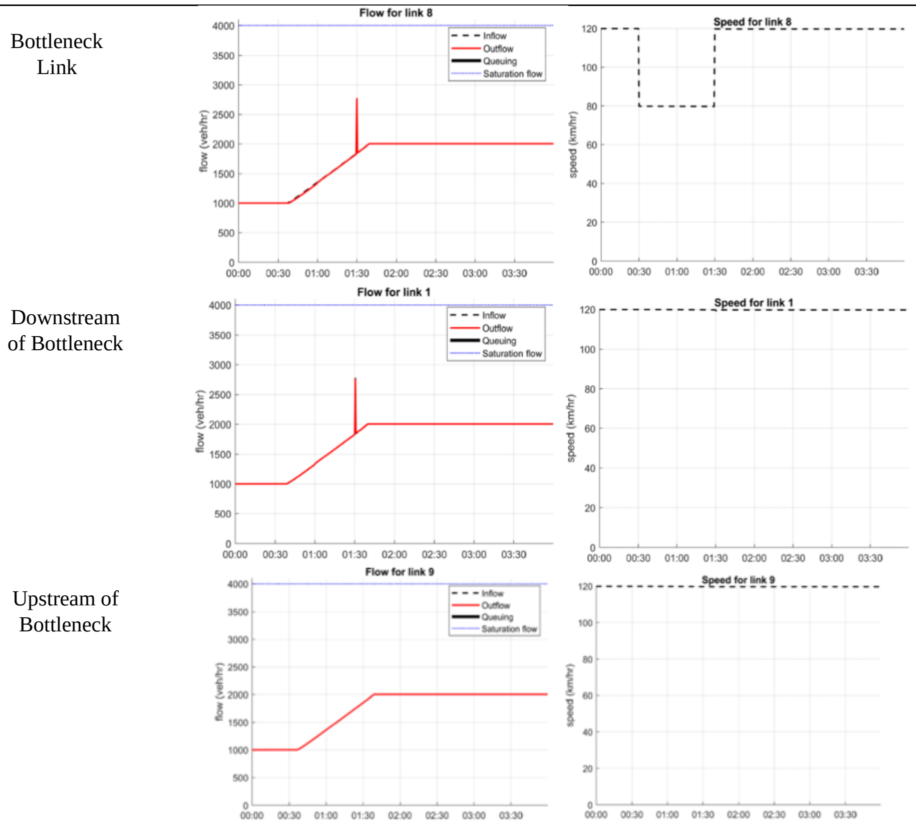
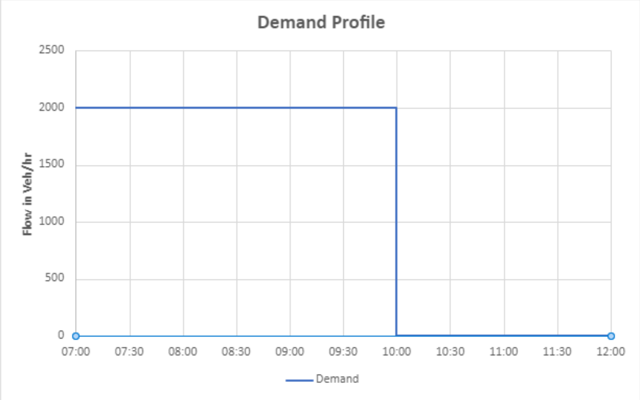
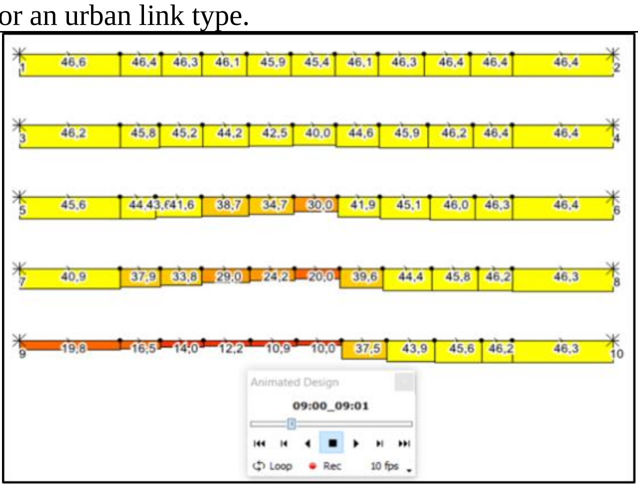
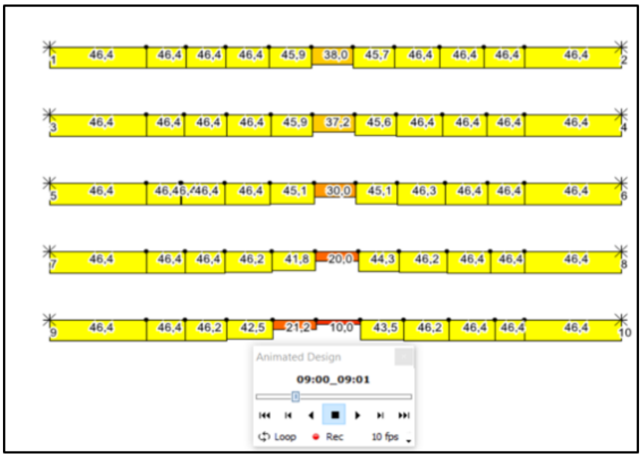
Figure 67: Speed values on the corridor links_1.6.1_StreamLine: MaDAM


Figure 68: Flow values on the corridor links_1.6.1_StreamLine: MaDAM
</td>
</tr>
</table>

	MaDAM offers a quite flexible facility to implement time-varying supply characteristics in the form of time varying fundamental diagrams, herewith enabling to model dynamic traffic control, dynamic behavior including VSL etc. We do observe a drop in speed when the control is activated at 07:30 as shown in Figure 67. However, the duration of the drop is extended beyond the deactivation of the control, which is not desirable. The speed also drops for the links downstream of the control link as a consequence of the convection term. However, the anticipation term does not account for a speed drop for the upstream links. Furthermore, as shown in the Figure 68, even the flows drop suddenly for the control links and the links downstream.
Score in EMMa	**2**
Results - MARPLE	 *Figure 69: Speed and Flow values on the corridor links_1.6.1_MARPLE* Results are in line with the expectation for the control link as speed drop occurs when the VSL is activated as shown in Figure 69. However, a sudden peak in flows (a value more than demand) is observed for the bottleneck link and downstream link, when the control is deactivated which is rather absurd. Furthermore, a speed drop is not observed in the downstream links, which may be due to the lack of second-order effects.
Score in EMMa	**3**
Results-StreamLine: eGLTM	Test not performed. Please refer to results of eGLTM in Test 1.5.1.
Score in EMMa	**0**
Test ID	**1.7.1**
Test Network	Refer to Figure 9
Test Description	The objective of the test is to understand the influence of urban and non-urban links within the same network. The variations in propagation behavior is checked through the same. The link

<table>
<tr><td></td><td>type of the central link (Link nr 6) is varied to understand the sensitivity of the urban and non-urban behavior.</td></tr>
<tr><td>MoP Evaluated -Link to EMMa</td><td>Fluctuation of traffic states over a series of urban and non-urban links.</td></tr>
<tr><td>Supply Properties</td><td>
<p>Table 14: Network properties for Test No. 1.7.1</p>
<table>
<tr><th>Supply</th><th>Link Type</th><th>O-D</th><th>No of Lane</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th></tr>
<tr><td>Network 1</td><td>Non-Central</td><td>1-2</td><td>1</td><td>0.3</td><td>50</td><td>30</td></tr>
<tr><td>Network 2</td><td>Non-Central</td><td>3-4</td><td>1</td><td>0.3</td><td>50</td><td>30</td></tr>
<tr><td>Network 3</td><td>Non-Central</td><td>5-6</td><td>1</td><td>0.3</td><td>50</td><td>30</td></tr>
<tr><td>Network 4</td><td>Non-Central</td><td>7-8</td><td>1</td><td>0.3</td><td>50</td><td>30</td></tr>
<tr><td>Network 5</td><td>Non-Central</td><td>9-10</td><td>1</td><td>0.3</td><td>50</td><td>30</td></tr>
</table>
<table>
<tr><th>Supply</th><th>Link Type</th><th>O-D</th><th>No of Lane</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th></tr>
<tr><td>Network 1</td><td>Central Link No 6</td><td>1-2</td><td>1</td><td>0.02</td><td>50</td><td>8</td></tr>
<tr><td>Network 2</td><td>Central Link No 6</td><td>3-4</td><td>1</td><td>0.02</td><td>40</td><td>8</td></tr>
<tr><td>Network 3</td><td>Central Link No 6</td><td>5-6</td><td>1</td><td>0.02</td><td>30</td><td>8</td></tr>
<tr><td>Network 4</td><td>Central Link No 6</td><td>7-8</td><td>1</td><td>0.02</td><td>20</td><td>8</td></tr>
<tr><td>Network 5</td><td>Central Link No 6</td><td>9-10</td><td>1</td><td>0.02</td><td>10</td><td>8</td></tr>
</table>
<p>The free-flow speeds in all links except Link No.6 is fixed at 50 Kmph. Over a series of 5 networks, the free-flow speed for Link No.6 is reduced from 50 Kmph to 10 Kmph. Length of all links except Link No. 6 is 300m. Length of Link No. 6 is 20m.</p>
</td></tr>
<tr><td>Demand Properties</td><td>


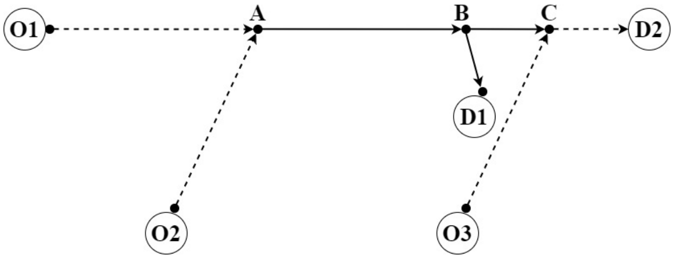
<p>Figure 70: Demand profile for Test No 1.7.1</p>
</td></tr>
</table>

	Demand between O and D is simulated in undersaturated conditions for all the links.
Expectation	The expectation here would be that an urban driver would undergo speed fluctuations more suddenly compared to a non-urban/motorway driver. This is under the assumption that urban drivers are more aggressive. The corresponding speed variations in the links upstream is also expected be sudden for an urban link type.
Results - StreamLine: MaDAM	 *Figure 71: Messmer & Papageorgiou (Non-Urban)_Flow propagation over the corridor links_1.7.1_StreamLine: MaDAM*  *Figure 72: Urban antipation term_Flow propagation over the corridor links_1.7.1_StreamLine: MaDAM* As can be observed from the propagation chart, the color and number suggest that StreamLine-MaDAM is able to deal with motorway (Figure 71) and Urban links (Figure 72) as expected by altering the anticipation term used, as stated in the figure captions respectively. The Highway anticipation terms (Figure 71), predicts the decrease in speed in the upstream links 1.5kms from the central link nr 6 (Free speed - 10Kmph, 20m link length) in case of the network nr 5(centroids 9 and 10). This falls in line with the assumption that in a highway stretch, the drivers anticipate a speed variation way ahead the disruption (smaller speed in this case). Similar trends are observed in the other 4 networks, with variation of speeds of the upstream links consistent with the defined free speed of the central link. In an urban setting (Figure 72), the speed variation more sudden and this is evident by means of the decrease in link speeds at 300m of distance from the central link as shown in propagation of the network nr 5 (centroids 9 and 10). An aggressive driver is expected here, where speed variation is more sudden than in case of a highway. Similar trends are observed in networks 1-4.
Score in EMMa	**4**

Results - MARPLE	MARPLE does not offer the option to vary the link types based on an urban and non-urban situation, other than varying the supply properties of the links externally.
Score in EMMa	**0**
Results-StreamLine: eGLTM	StreamLine does not offer the option to vary the link types based on an urban and non-urban situation, other than varying the supply properties of the links externally.
Score in EMMa	**0**

2. Node model – merge & diverge

The second series of testing is aimed at checking the node model capabilities of the DTA models. Essentially the criteria laid out by (Tampère, Corthout, Cattrysse, & Immers, 2011) as described in Section 2.3.2.4 will be checked in these tests. The test networks mainly adopted from (FakhraeiRoudsari, Huang, & Tampère, 2015).

<table>
<tr><th>Test ID</th><th>2.1.1</th></tr>
<tr><td>Test Network</td><td>
</td></tr>
<tr><td>Test Description</td><td>The aim of this test is to check single merge behavior where the receiving link’s capacity is the constraint and both inflow of sending links exceed their reduced outflow capacity.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-merge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>
<i>Table 15: Network properties for Test No. 2.1.1</i>
<table>
<tr><th></th><th>O1A</th><th>O2A</th><th>AB</th><th>BC</th><th>BD1</th><th>O3C</th><th>CD2</th></tr>
<tr><td>Length in Kms</td><td>5</td><td>5</td><td>5</td><td>1</td><td>1</td><td>5</td><td>1</td></tr>
</table>
<table>
<tr><th>Supply</th><th>Type</th><th>Capacity</th><th>Length (Km)</th><th>Free Flow Speed</th><th>Speed at Cap</th><th>No of Lanes</th></tr>
<tr><td>Highway Links</td><td>5</td><td>1000</td><td>0.3</td><td>120</td><td>90</td><td>1</td></tr>
<tr><td>On-Ramp Links</td><td>10</td><td>1000</td><td>0.3</td><td>80</td><td>30</td><td>1</td></tr>
<tr><td>Off-Ramp Links</td><td>10</td><td>1000</td><td>0.3</td><td>80</td><td>30</td><td>1</td></tr>
<tr><td>Connector</td><td>1</td><td>3000</td><td>1</td><td>80</td><td>30</td><td>1</td></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td>Demand Properties</td>
<td>

Table 16: Demand Matrices for Test No 2.1.1

<table>
<tr><th></th><th>from</th><th>to</th><th></th><th>from</th><th>to</th><th></th><th>from</th><th>to</th><th></th><th>from</th><th>to</th></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>0</td><td>800</td><td>O1</td><td>0</td><td>800</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>100</td><td>O3</td><td>0</td><td>600</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table>

</td>
</tr>
<tr>
<td>Expectation</td>
<td>Capacity reduction can be expected at merge node C, to restrict the outflow values of the sending links, to maintain continuity in link CD2.</td>
</tr>
<tr>
<td>Results - StreamLine: MaDAM</td>
<td>

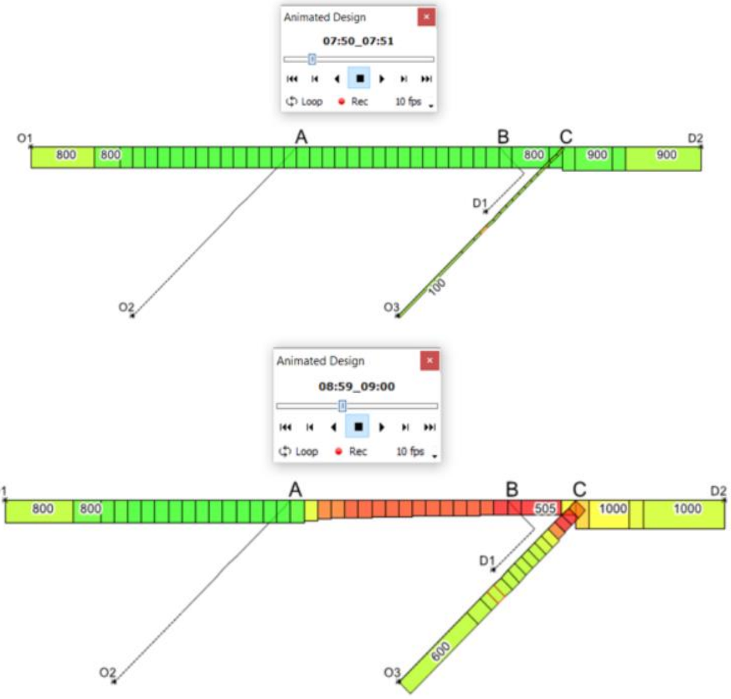


Figure 73: Flow propagation chart for the highway corridor_2.1.1_StreamLine: MaDAM

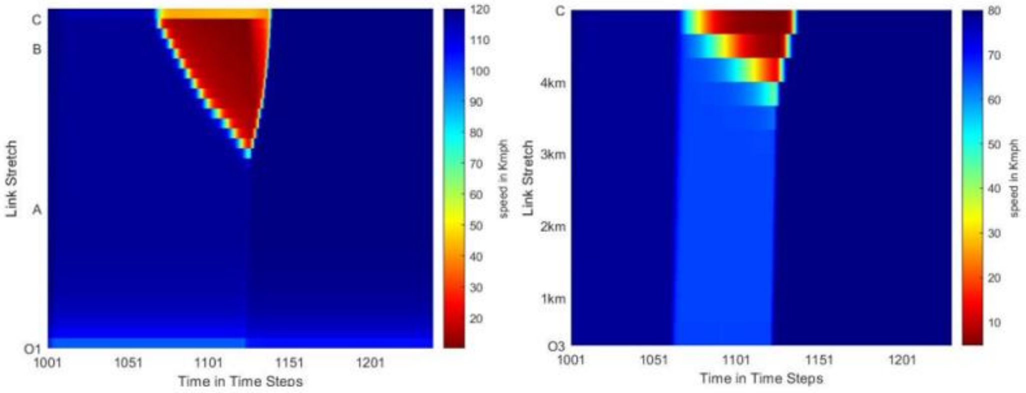
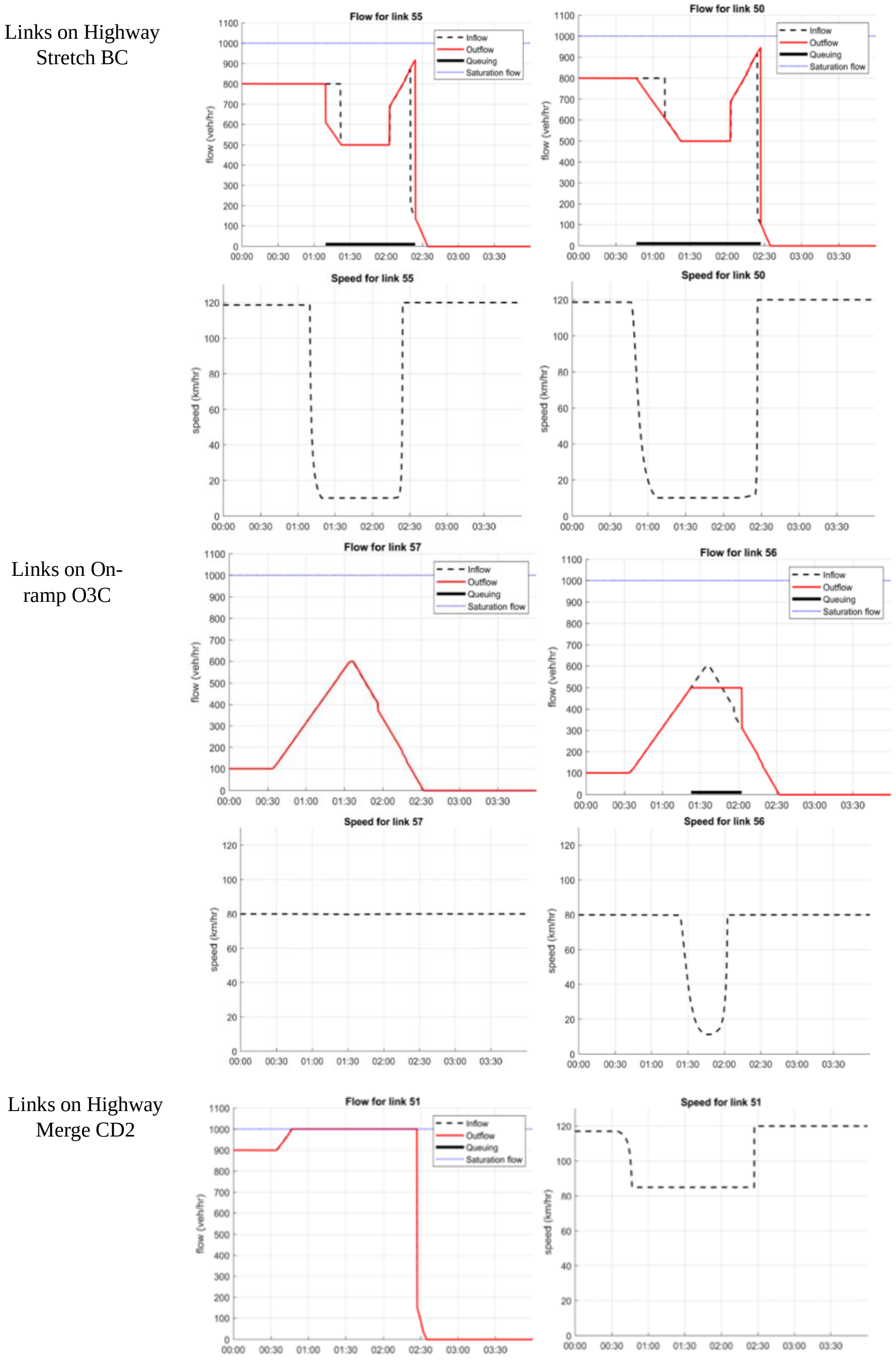


Figure 74: Space Time Diagram of Highway Stretch (Left) and On-ramp stretch O3C(Right)_2.1.1_StreamLine: MaDAM

From the propagation diagram in Figure 73, it can be viewed that when the outgoing link capacity is 1000 veh/hr, the capacity of the incoming links is reduced to nearly half (capacity proportionality), thereby ensuring the continuity principle. The colors in the propagation chart represent the speeds varying between 0-120 kmph, red-green respectively. Consequently, this has resulted in congestion in both the sending links (Demand > Capacity) and thereby spillback

</td>
</tr>
</table>

<table>
<tr><td></td><td>upstream. Interestingly, the second order effects of traffic such capacity drop and speed term modifications are not seen in the model results. It can be interpreted that StreamLine over-rules the second order effects while incorporating the node constraints. The congestion patterns are described in the X-T Diagram in Figure 74.</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
<tr><td>Results - MARPLE</td><td>


<i>Figure 75: Flows and speeds on the links under study_2.1.1_MARPLE</i>

The issue of short links was posed while running the model. The time step for propagation is reduced to 8sec needs to avoid this. The merge node behavior of capacity restriction is satisfactorily modelled. With the capacity-->flow of the sending links restricted to that of the receiving link (continuity principle). This becomes the case when the demand > capacity of the
</td></tr>
</table>

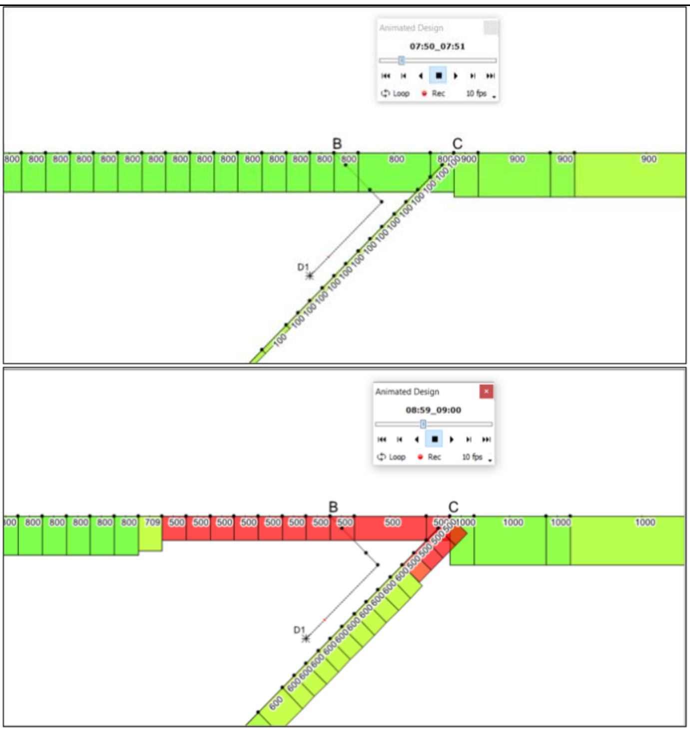
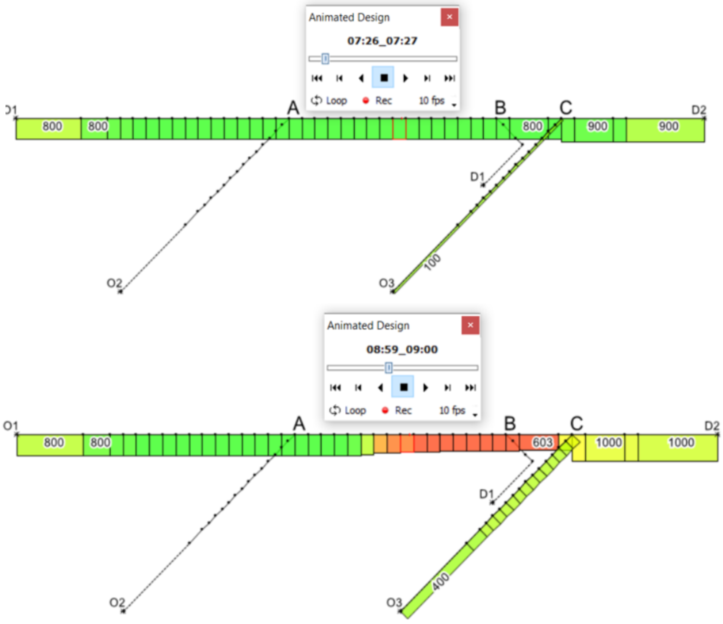
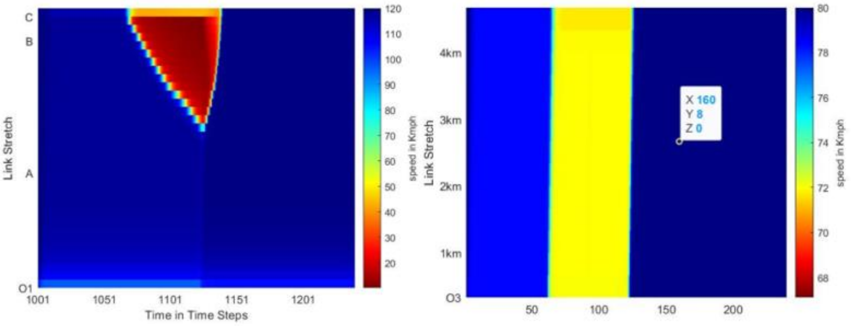
<table>
<tr><td></td><td>merge highway link. Consequently, queuing is observed in both the sending road sections as the demand in the links exceed their reduced outflow capacity. Additionally, while inspecting the propagation of the second on-ramp road stretch O3C, the spillback is found missing beyond the link directly upstream of the merge node (Link nr 56).</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
<tr><td>Results-StreamLine: eGLTM</td><td>


Figure 76:Flow propagation chart for the highway corridor_2.1.1_StreamLine: eGLTM

Results are in line with the expectation. From the propagation diagram, it can be viewed that when the outgoing link capacity is 1000 veh/hr, the capacity of the incoming links is reduced to exactly half, thereby ensuring the continuity principle. As expected, the effect of capacity reduction in the links upstream due to spillback is lesser in case of eGLTM compared to MaDAM due to the lack of second order effects.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Test ID</td><td>2.1.2</td></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>The aim of this test is to check single merge behavior where the receiving link’s capacity is the constraint and only highway’s inflow exceed its constrained outflow capacity.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-merge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1</td></tr>
<tr><td>Demand Properties</td><td>

Table 17: Demand Matrices for Test No. 2.1.2

<table>
<tr><td></td><td>from</td><td>to</td><td></td><td>from</td><td>to</td><td></td><td>From</td><td>to</td><td></td><td>from</td><td>To</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>0</td><td>800</td><td>O1</td><td>0</td><td>800</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>100</td><td>O3</td><td>0</td><td>400</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table>
</td></tr>
</table>

Expectation	Congestion is expected only in the highway stretch as a consequence of capacity reduction in the sending links in the highway.
Results - StreamLine: MaDAM	 *Figure 77: Flow propagation chart for the highway corridor_2.1.2_StreamLine: MaDAM*  *Figure 78:Space Time Diagram of Highway Stretch (Left) and On-ramp stretch O3C(Right)_2.1.2_StreamLine: MaDAM* From the propagation diagram in Figure 77, it can be viewed that when the outgoing link capacity is 1000 veh/hr, the capacity of the incoming links are reduced to half, thereby ensuring the continuity principle. The highway merge link has obtained the excess of 100 veh/hr capacity due to the smaller capacity of the on-ramp link, thereby ensuring continuity as the ouflow of the on-ramp link is only 400 veh/hr. Interestingly, the second order effects of traffic such capacity drop and speed term modifications are not seen in the model. It can be interpreted that SL over-rules the second order effects by incorporating the node constraints. The congestion patterns are described in the X-T Diagram shown in Figure 78.
Score in EMMa	**3**

<table>
<tr>
<td>Results - MARPLE</td>
<td>

Links on Highway Stretch BC

Links on On-ramp O3C

Links on Highway Merge CD2

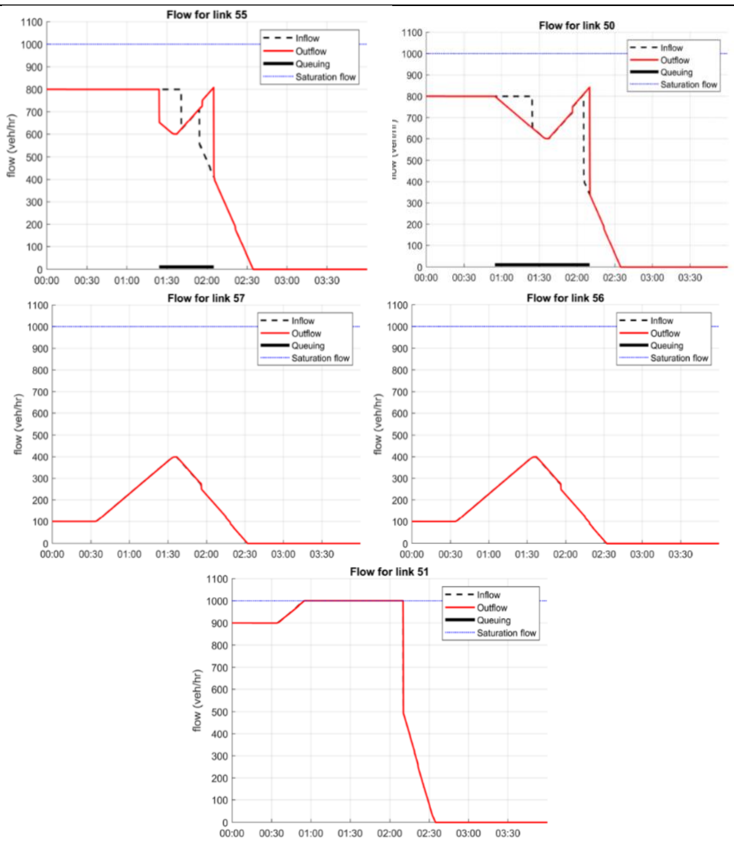


Figure 79: Flow propagation values in links of the highway corridor_2.1.2_MARPLE

Ignoring the issued faced in propagation already described in Test series 1, the merge behavior was indeed modelled as per theoretical expectations as shown. The on-ramp links have an incoming flow of 400 veh/hr, which is less than the reduced capacity of 500 veh/hr. The highway link in BC, receive this additional 100 veh/h. Therefore the outflow becomes (500-400)+500 = 600 veh/hr, which can be observed through the flow propagation charts. A queue is absent in the O3C link, as expected as the demand here is less than the capacity.

</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>3</td>
</tr>
<tr>
<td>Results-StreamLine: eGLTM</td>
<td>

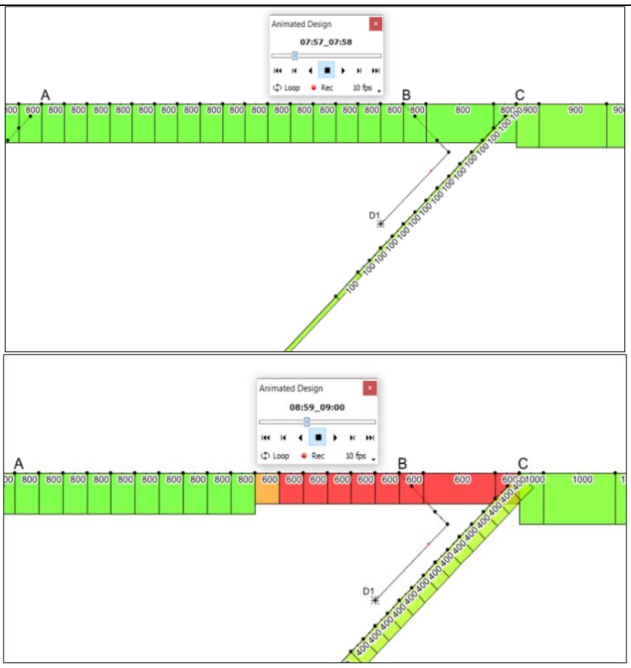
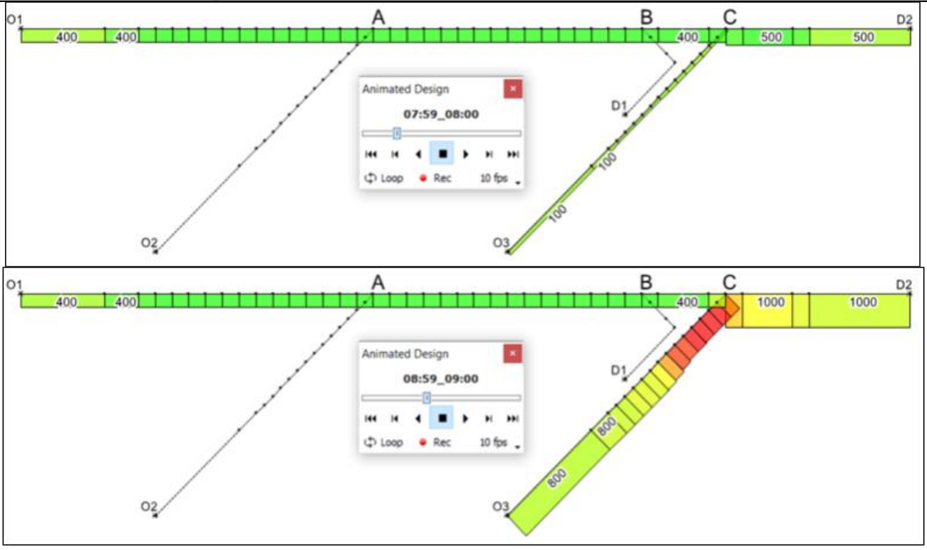


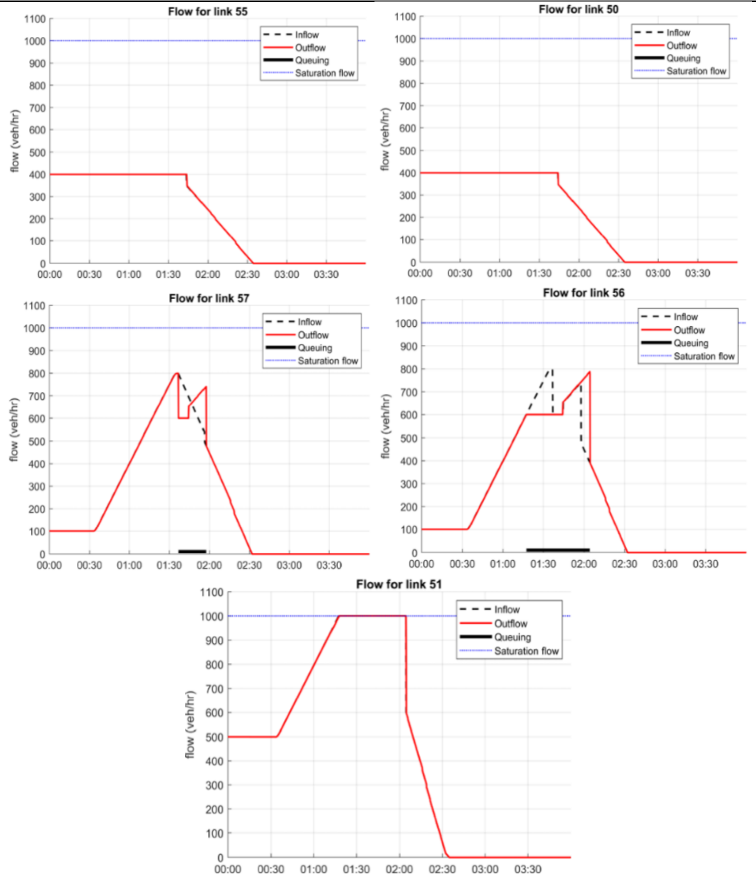
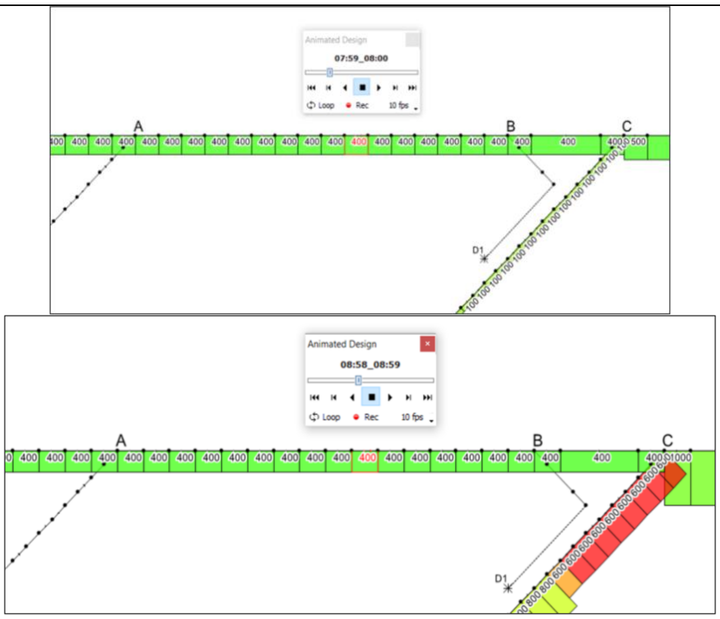
Figure 80: Flow propagation chart for the highway corridor_2.1.2_StreamLine: eGLTM

</td>
</tr>
</table>

<table>
<tr><td></td><td>As per expectation the capacity reduction in the sending links, have reduced the outflow to 500 veh/hr thereby satisfying the merge node constraints. As the inflow from the merge link in On-ramp stretch O2C is lesser than the capacity there wouldn’t be any queue or congestion as shown in Figure 80Figure 77. The queue is seen in links in highway stretch BC, which has an inflow rate greater than the reduced capacity by about 300 veh/hr (800-500) veh/hr. The highway merge link has obtained the excess of 100 veh/hr capacity due to the smaller capacity of the on-ramp link, thereby ensuring continuity as the outflow of the on-ramp link is only 400 veh/hr.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Test ID</td><td>2.1.3</td></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>The aim of this test is to check single merge behavior where the receiving link’s capacity is the constraint and only on-ramp’s inflow exceed its constrained outflow capacity.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-merge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1</td></tr>
<tr><td>Demand Properties</td><td><i>Table 18: Demand Matrices for Test No. 2.1.2</i>
<table>
<tr><td></td><td>from</td><td>to</td><td></td><td>from</td><td>to</td><td></td><td>From</td><td>to</td><td></td><td>from</td><td>To</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>0</td><td>400</td><td>O1</td><td>0</td><td>400</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>100</td><td>O3</td><td>0</td><td>800</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>Congestion is expected only in the on-ramp stretch as a consequence of capacity reduction in the sending links in the on-ramp.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

<i>Figure 81: Flow propagation chart for the highway corridor_2.1.3_StreamLine: MaDAM</i>

Results are in line with the expectation as shown in Figure 81. The queue formation in this case occurs in the On-ramp link O3C. There is no queue on the highway, as inflow<capacity in this stretch. Congestion pattern is further evident in the X-T diagrams.</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
</table>

Results - MARPLE	Links on Highway Stretch BC Links on On-ramp O3C Links on Highway Merge CD2  *Figure 82: Flow propagation values in links of the highway corridor_2.1.3_MARPLE* Results in the link loads are consistent with the expectation. In the highway merge link BC, the flow values are less than the reduced capacity from the merge link. The additional capacity is used by the converge link in the on-ramp O3C, which portrays a load of 600 veh/hr. However, the issues present in the propagation for MARPLE is still visible in the node model. The extent of spillback in the on-ramp stretch is lesser than expectation. Overall, the node model simulates the results fairly consistent with the expectations.
Score in EMMa	**3**
Results-StreamLine: eGLTM	 *Figure 83: Flow propagation chart for the highway corridor_2.1.3_StreamLine: eGLTM*

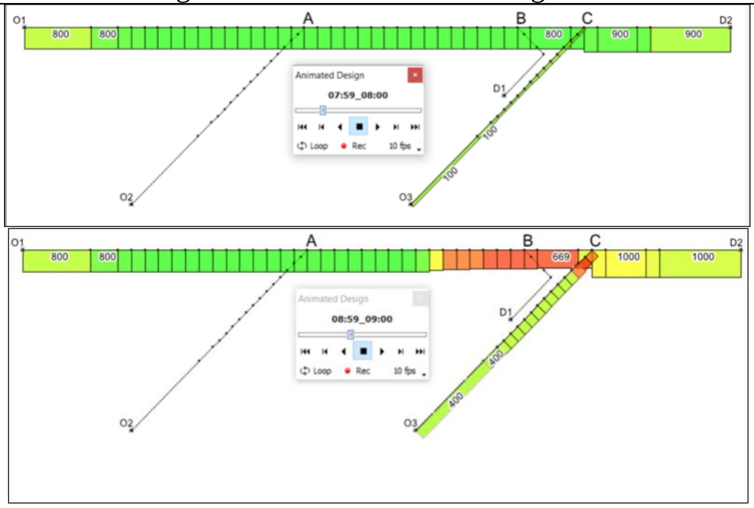
<table>
<tr><td></td><td>Expectation and simulation results are similar to previous case. Just that the queue formed here is on the On-ramp link O3C. The splitting rates at the merge node also seems to divide the outflows precisely into the sending links as shown in Figure 83. There is no queue on the highway, as inflow<capacity in this stretch. Congestion pattern is further evident in the X-T diagrams.</td></tr>
<tr><td>Score in EMMa</td><td>**4**</td></tr>
<tr><td>**Test ID**</td><td>**2.1.4**</td></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>The aim of this test is to check single merge behavior where the receiving link's capacity is the constraint and capacity of the highway merge link is twice more than that of the on-ramp link.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td>*Node model-merge behavior*</td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1, except the links on Highway Merge Stretch BC have capacity = 2000 veh/hr</td></tr>
<tr><td>Demand Properties</td><td>*Table 19: Demand Matrices for Test No. 2.1.4*
<table>
<tr><td></td><td>**from**</td><td>**to**</td><td></td><td>**from**</td><td>**to**</td><td></td><td>**From**</td><td>**to**</td><td></td><td>**from**</td><td>**To**</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>**D1**</td><td>**D2**</td><td></td><td>**D1**</td><td>**D2**</td><td></td><td>**D1**</td><td>**D2**</td><td></td><td>**D1**</td><td>**D2**</td></tr>
<tr><td>**O1**</td><td>0</td><td>800</td><td>**O1**</td><td>0</td><td>800</td><td>**O1**</td><td>0</td><td>0</td><td>**O1**</td><td>0</td><td>0</td></tr>
<tr><td>**O3**</td><td>0</td><td>100</td><td>**O3**</td><td>0</td><td>400</td><td>**O3**</td><td>0</td><td>0</td><td>**O3**</td><td>0</td><td>0</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>Congestion is expected both the highway stretch and the on-ramp assuming capacity proportionality while calculating outflow values of the sending links.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

Figure 84: Flow propagation chart for the highway corridor_2.1.4_StreamLine: MaDAM

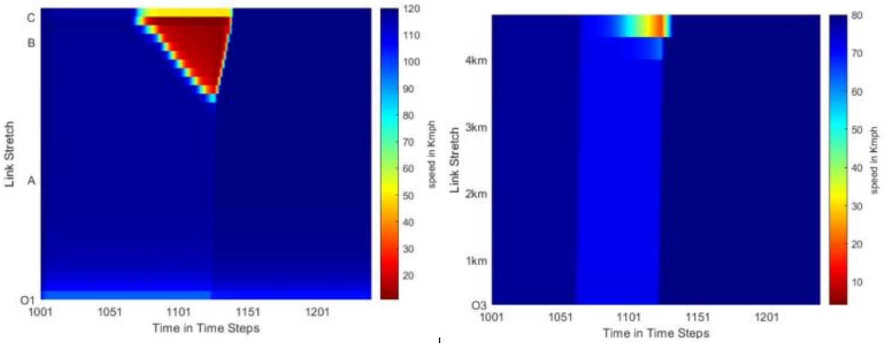
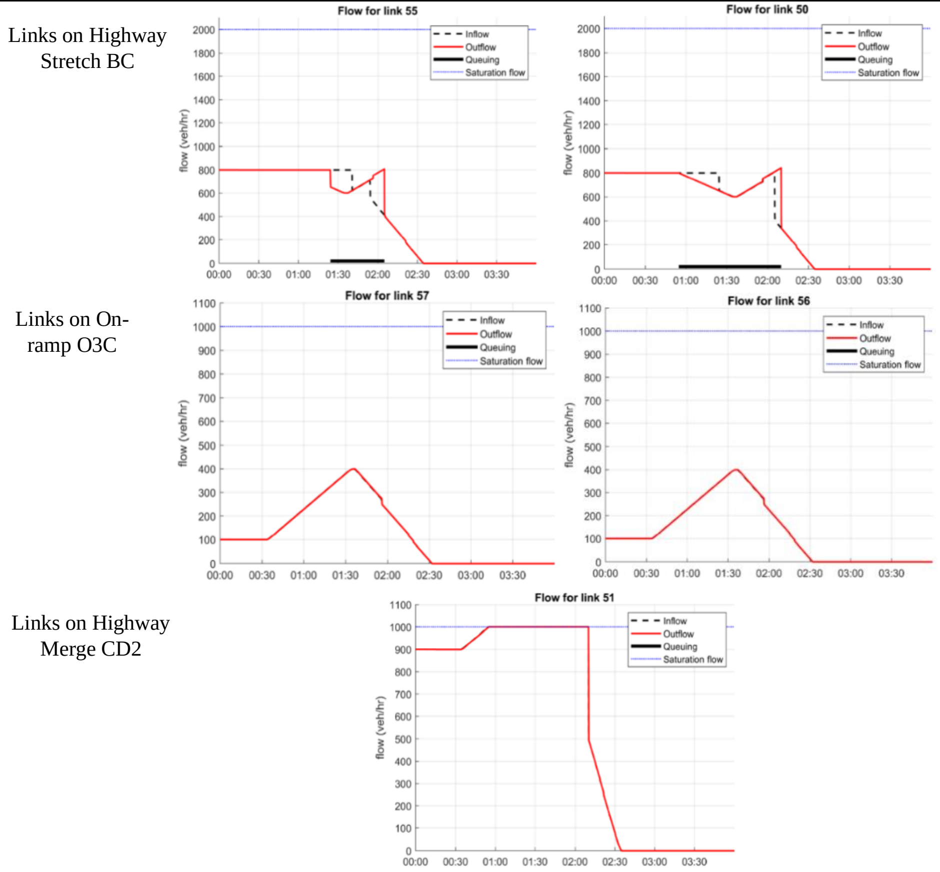
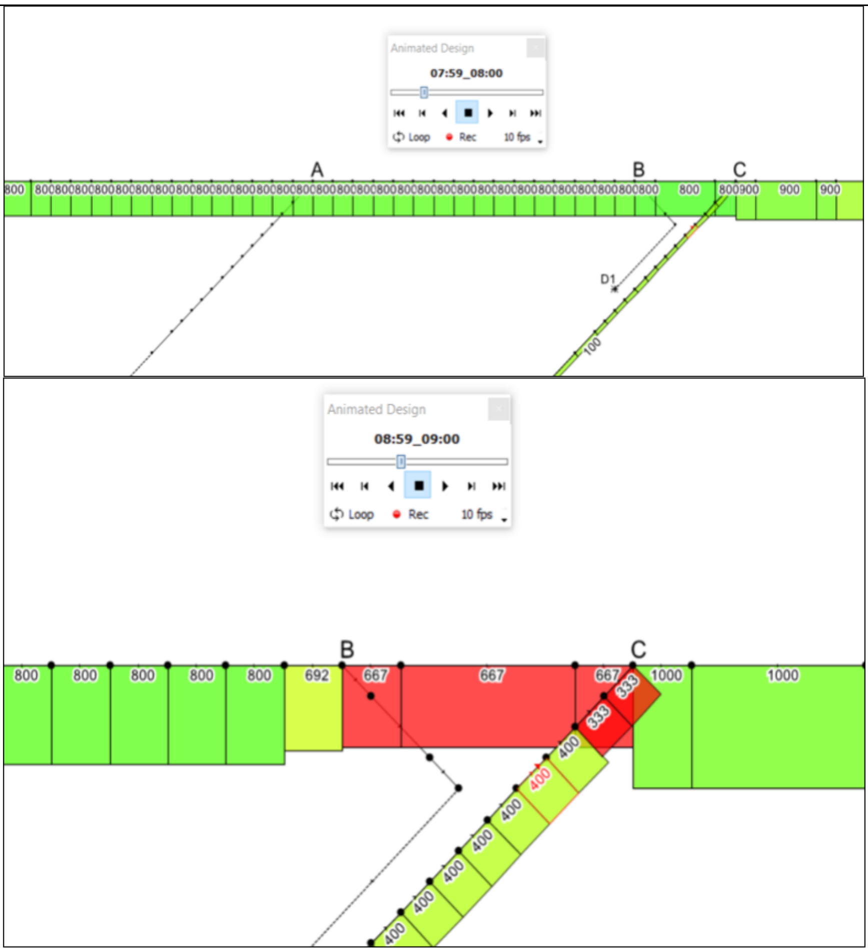
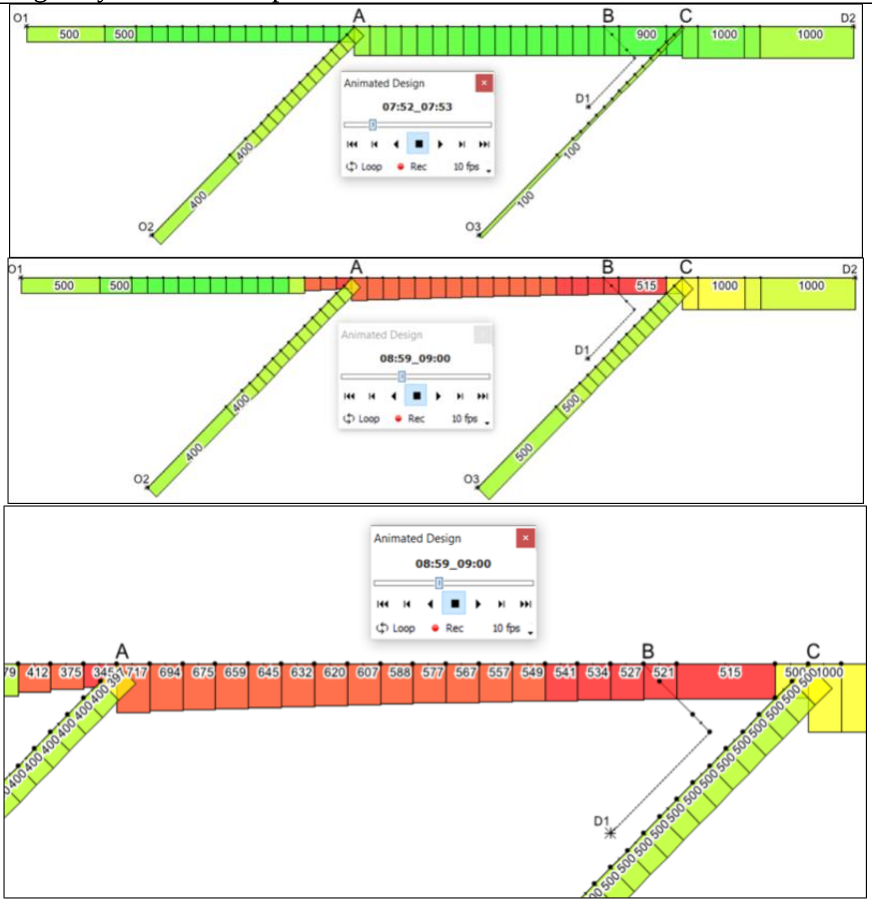
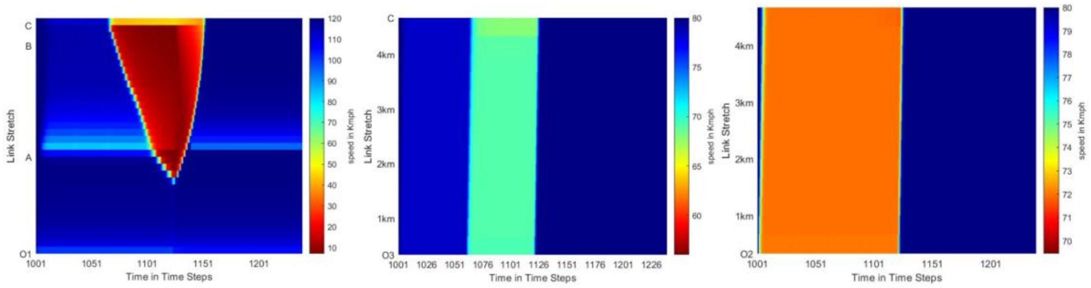


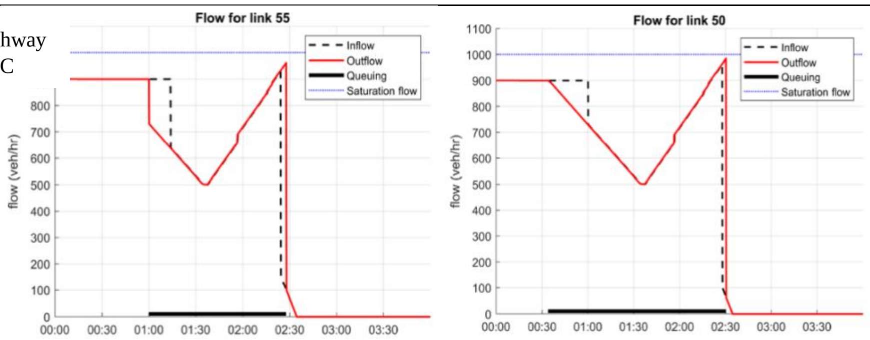
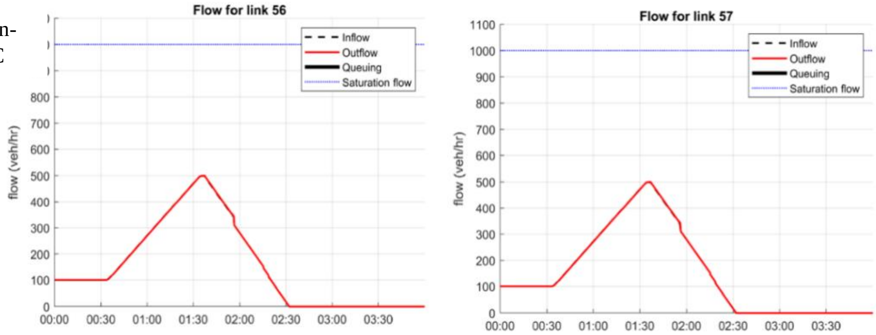
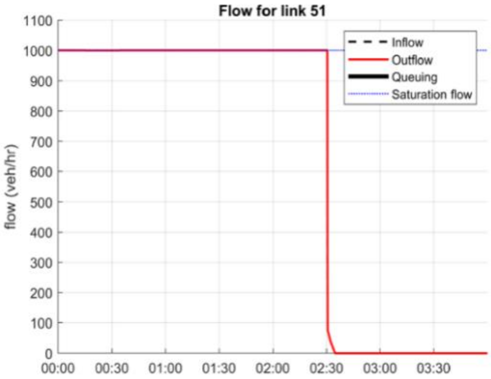
Figure 85: Space Time Diagram of Highway Stretch (Left) and On-ramp stretch O3C(Right)_2.1.4_StreamLine: MaDAM</td></tr>
</table>

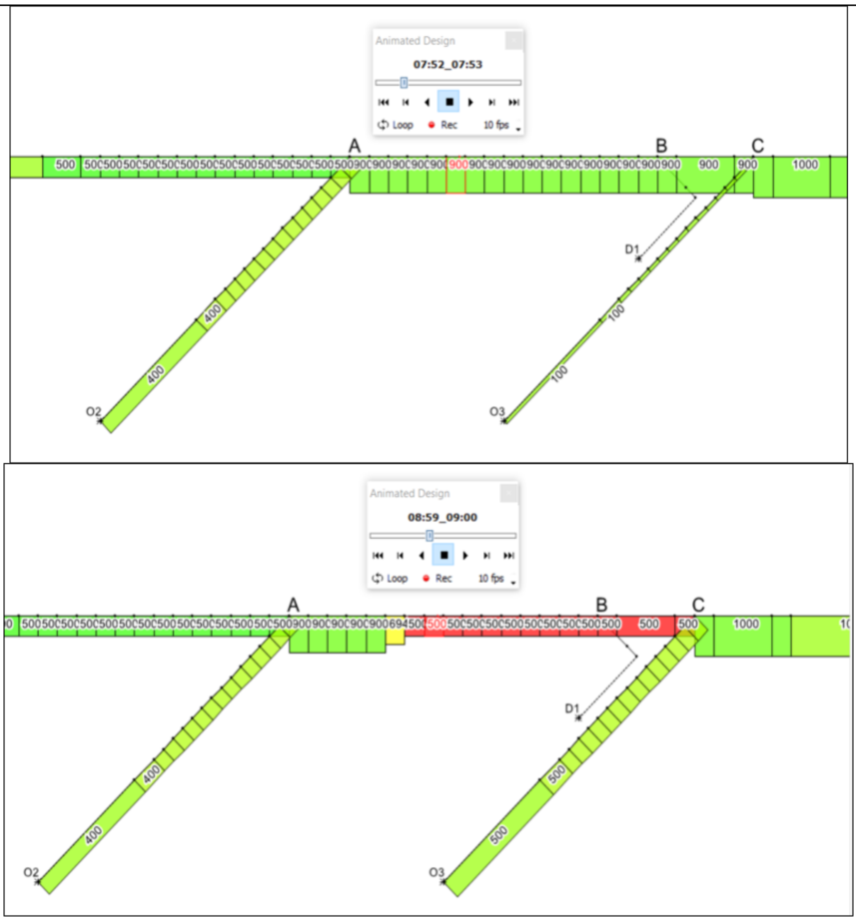
	As expected, congestion is present in both the highway and the on-ramp link as shown in Figure 84. Further reinforcing the expectation is that the congestion in the highway link is approximately twice that of the on-ramp link O3C, as a direct consequence of capacity proportionality. The congestion pattern is evident in the X-T diagrams Figure 85.
Score in EMMa	**4**
Results - MARPLE	 *Figure 86: Flow propagation values in links of the highway corridor_2.1.4_MARPLE* Results are in line with the expectation. Traffic gets blocked at merge node C, as a result of the capacity constraint in the highway stretch CD2. The distribution of outflow values in the sending links is according to the number of lanes in the incoming sending links in MARPLE, as opposed to its capacity in StreamLine. The number of lanes is equal in all the links in the network. Therefore, as the demand in the on-ramp O3C, is less than the capacity, the balance is shifted to the highway merge link in BC, reducing the outflow of the merge link to 600 veh/hr, maintaining the principle of continuity. The results are as shown in Figure 86.
Score in EMMa	**4**

Results-StreamLine: eGLTM	 *Figure 87: Flow propagation chart for the highway corridor_2.1.4_StreamLine: eGLTM* As expected, congestion is present in both the highway and the on-ramp link. Further reinforcing the expectation is that the congestion in the highway link is approximately twice that of the on-ramp link O3C, as a direct consequence of capacity proportionality.
Score in EMMa	**4**
Test ID	**2.2.1**
Test Network	Same as 2.1.1
Test Description	To understand merging behavior in the event of congestion, triggered by spillback from a more downstream bottleneck.
MoP Evaluated - Link to EMMa	*Node model-merge behavior*
Supply Properties	Same as 2.1.1
Demand Properties	*Table 20: Demand Matrices for Test No. 2.2.1*

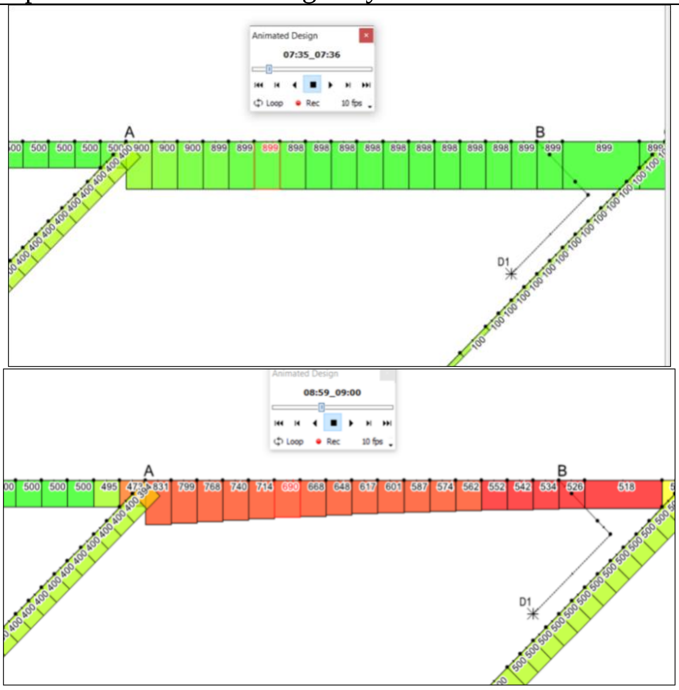
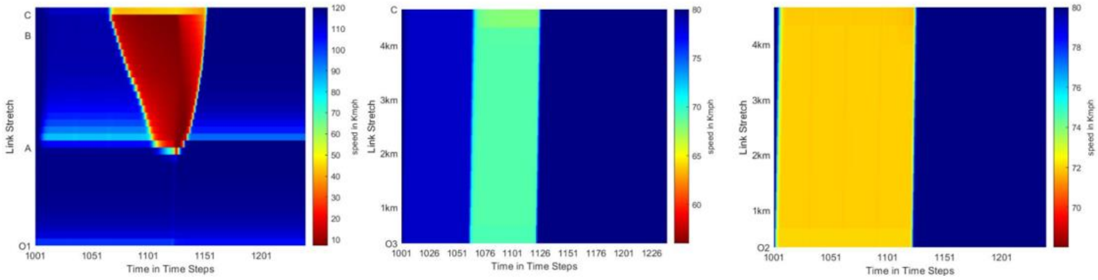
	from	**to**		**from**	**to**		**From**	**to**		**from**	**To**
	07:00	08:00		08:00	09:00		09:00	10:00		09:00	10:00
	D1	**D2**		**D1**	**D2**		**D1**	**D2**		**D1**	**D2**
O1	0	500	**O1**	0	500	**O1**	0	0	**O1**	0	0
O2	0	400	**O2**	0	400	**O2**	0	0	**O2**	0	0
O3	0	100	**O3**	0	500	**O3**	0	0	**O3**	0	0

Expectation	Congestion is expected in the highway stretch CD2, as a consequence of oversaturation. The congestion is expected to trigger a spillback upstream, reducing the links speeds of vehicles entering the highway from On ramp sections as well.
Results - StreamLine: MaDAM	 *Figure 88: Flow propagation chart for the highway corridor_2.2.1_StreamLine: MaDAM*  *Figure 89: Space Time Diagram of Highway Stretch (Left) and On-ramp stretch O3C(middle) and on-ramp stretch O2A(Right)_2.2.1_StreamLine: MaDAM* The results match the expectation that the segment from node C towards centroid D2 is congested and it spills back into the the subsequent upstream links. Furthermore, to maintain continuity the merge links in the stretch BC and O3C gets a reduced capacity of 500 veh/hr. This is further expected to reduced 250 veh/hr in the stretch O1A and O2A with capacity values of 250 veh/hr each. As can be observed from the figures, the capacity reduction in the links is not strictly followed. The heuristics in the model accounts for the upstream spillback overrules the capacity reduction constraint and maintains the flow higher than the reduced capacity upstream from the congested node. This could be due to the absence of an explicit node model. The congestion pattern is as shown in Figure 89.
Score in EMMa	2

Results - MARPLE	Links on Highway Stretch BC  Links on On-ramp O3C  Links on Highway Merge CD2  *Figure 90: Flow propagation values in links of the highway corridor_2.2.1_MARPLE* The expectation would be that during the second hour of the simulation, a bottleneck due to oversaturation in the highway merge link in CD2 will cause congestion in the links BC and O3C causing a reduced capacity to the tune of 500 veh/hr in both the links as a result of proportionality based on number of lanes. As a result of this, the reduced capacity in the upstream links is expected to cause a reduced capacity in the first highway merge links - O1A and O2A to the tune of 250 veh/hr. This would subsequently cause congestion in these links. However, this expectation is not met by the model results and capacity reduction is not caused in the links upstream from the bottleneck link. The absence of congestion is also reinforced by the speed distribution diagrams.
Score in EMMa	**2**

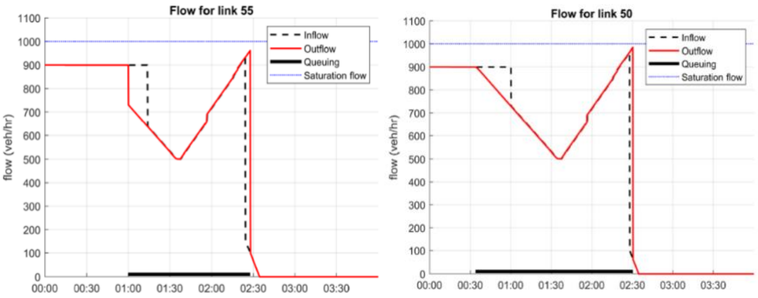
Results-StreamLine: eGLTM	 *Figure 91: Flow propagation chart for the highway corridor_2.2.1_StreamLine: eGLTM* Results meet the expectation that the segment from C towards D2 is congested and it spills back into the subsequent upstream links. Furthermore, to maintain the capacity proportionality, the merge links in the stretch BC and O3C gets a reduced capacity of 500 veh/hr.
Score in EMMa	**4**
Test ID	**2.2.2**
Test Network	Same as 2.1.1
Test Description	To understand merging behavior in the event of congestion, triggered by spillback from a more downstream bottleneck. Change with the previous case is to check the effect when the capacity of the sending links is different. To see the difference in proportionality.
MoP Evaluated - Link to EMMa	*Node model-merge behavior*
Supply Properties	Same as 2.1.1, the capacity of all the links in highway stretch O1A is changed to 5000 veh/hr.
Demand Properties	*Table 21: Demand Matrices for Test No. 2.2.1*

	from	to		from	to		From	to		from	To
	07:00	08:00		08:00	09:00		09:00	10:00		09:00	10:00
	D1	**D2**		**D1**	**D2**		**D1**	**D2**		**D1**	**D2**
O1	0	500	**O1**	0	500	**O1**	0	0	**O1**	0	0
O2	0	400	**O2**	0	400	**O2**	0	0	**O2**	0	0
O3	0	100	**O3**	0	500	**O3**	0	0	**O3**	0	0

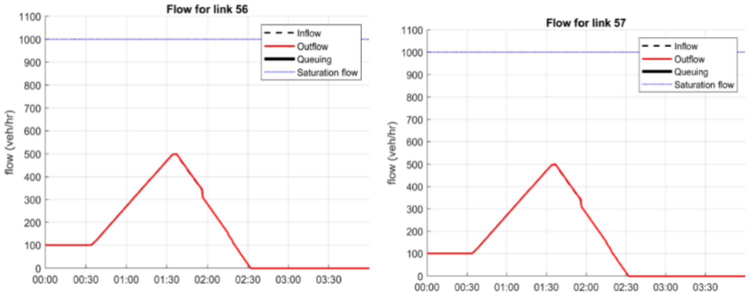
Expectation	Other than the expectation of a congestion caused in the highway stretch CD2 and subsequent spillback, the expectation would be that the nodal split at A, will show a capacity proportionality of 1/6th to the on-ramp O2A and 5/6 to the highway link O1A.
Results - StreamLine: MaDAM	 *Figure 92: Flow propagation chart for the highway corridor_2.2.2_StreamLine: MaDAM*  *Figure 93: Space Time Diagram of Highway Stretch (Left) and On-ramp stretch O3C(middle) and on-ramp stretch O2A(Right)_2.2.2_StreamLine: MaDAM* The results do not match the expectation. A separate test was also conducted by propagating the demand for one more hour from 09:00 to 10:00, during which the results were similar to Figure 92(not shown in this results section). Hence, we can conclude, the merge node property of capacity proportionality. The 2nd order heuristics in the model, overrules nodal constraints. The congestion pattern can be viewed through the X-T diagrams.
Score in EMMa	2

Results - MARPLE

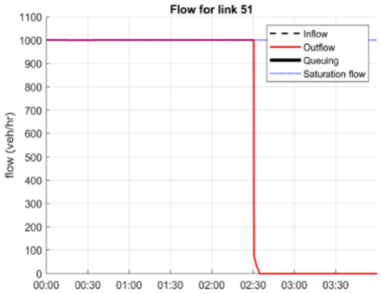
Links on Highway Stretch BC



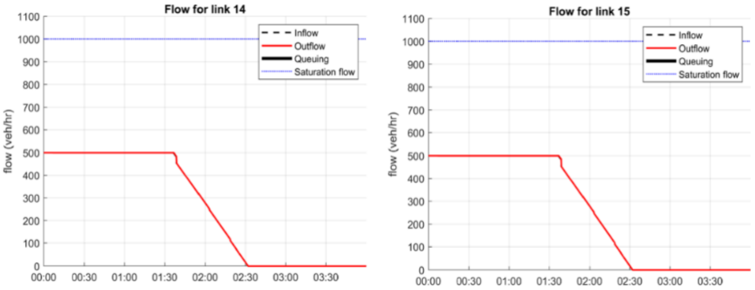
Links on On-ramp O3C



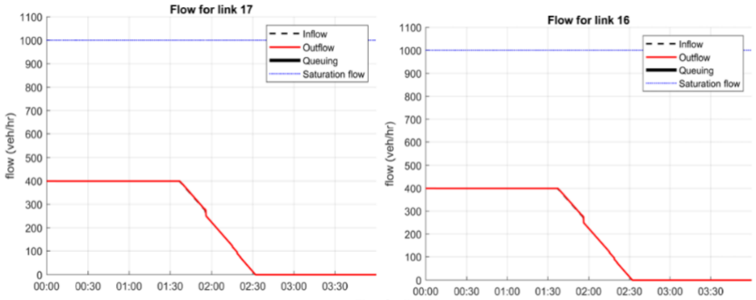
Links on Highway Merge CD2



Links on Highway Stretch O1A



Links on On-ramp O2A



Links on Highway Merge AB

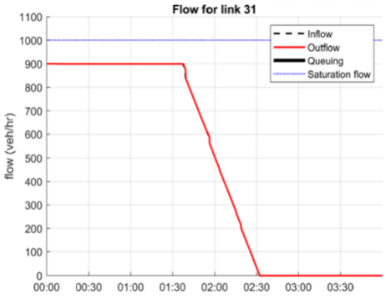
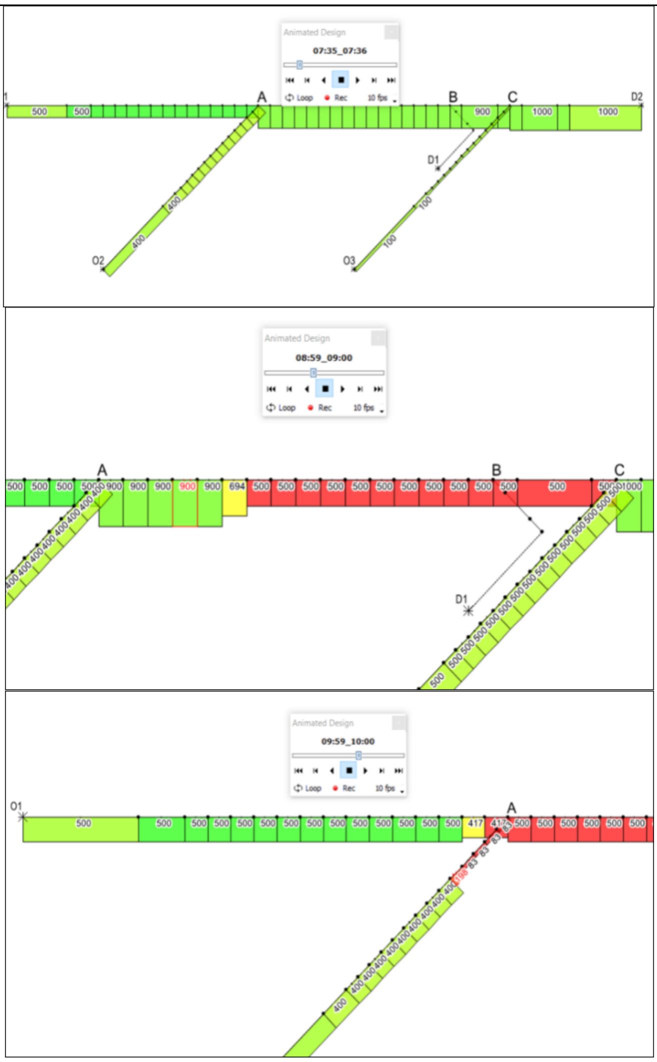


Figure 94: Flow propagation values in links of the highway corridor_2.2.2_MARPLE

<table>
<tr><td></td><td>The expectation would be that during the second hour of the simulation, a bottleneck due to oversaturation in the highway merge link in CD2 will cause congestion in the links BC and O3C causing a reduced capacity to the tune of 500 veh/hr in both the links assuming capacity proportionality. As a result of this, the reduced capacity in the upstream links is expected to cause a reduced capacity in the first highway merge links - O1A and O2A to the tune of 250 veh/hr. This would subsequently cause congestion in these links. We do observe an instance of congestion and queuing in the links directly upstream from CD2 i.e., BC but not links beyond the same. Thus expectation is not completely met by the model results and capacity reduction is not caused in the links upstream from the bottleneck link which is evident from the propagation chart in Figure 94.</td></tr>
<tr><td>Score in EMMa</td><td>2</td></tr>
<tr><td>Results-StreamLine: eGLTM</td><td>


<i>Figure 95: Flow propagation values in links of the highway corridor_2.2.2_eGLTM</i>

The expectation was that the nodal split at A, will show a capacity proportionality of 1/6th to the on-ramp O2A and 5/6 to the highway link O1A. Results lie completely in line with the expectation as shown in Figure 95. To further reinstate this expectation, the propagation was continued for one more hour demand from 09:00-10:00. At node A, it can be observed that the flow values in the links in highway stretch O1A are 417(5/6th of 500 - receiving flow) and the links in onramp O2A have a flow of 83 (1/6th of 500). Thus, capacity proportionality is strictly followed.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
</table>

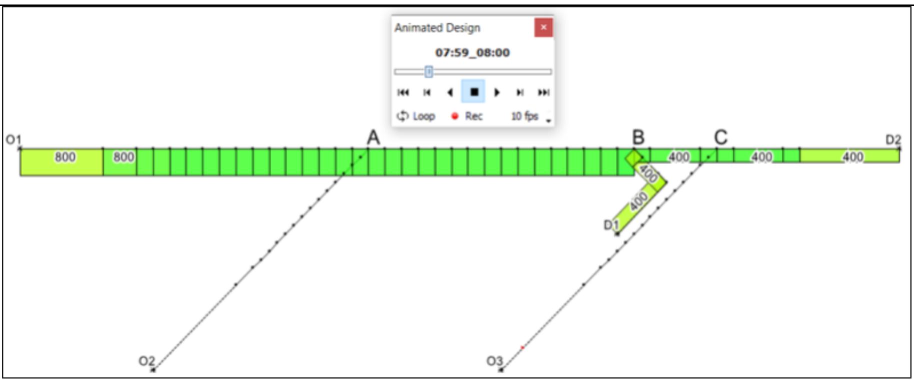
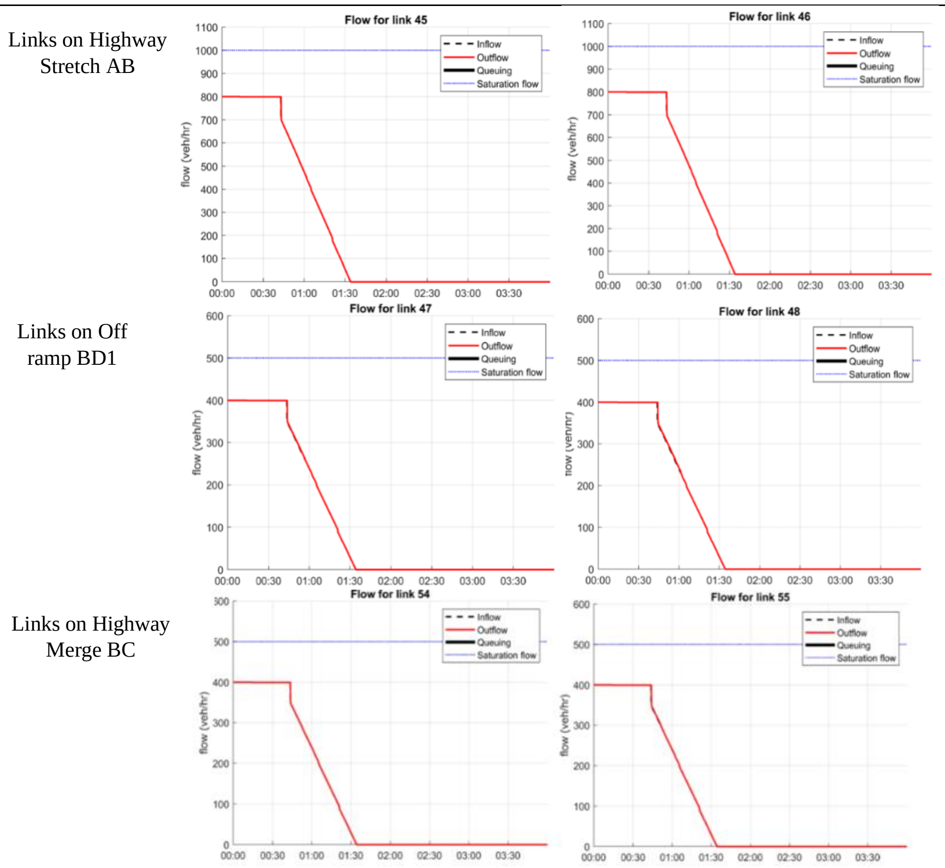
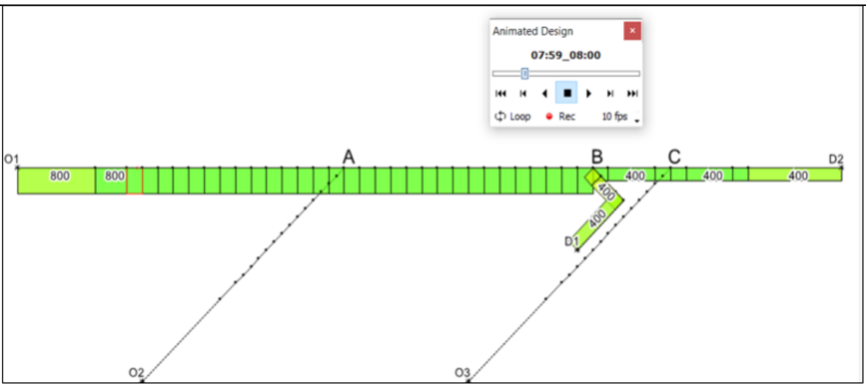
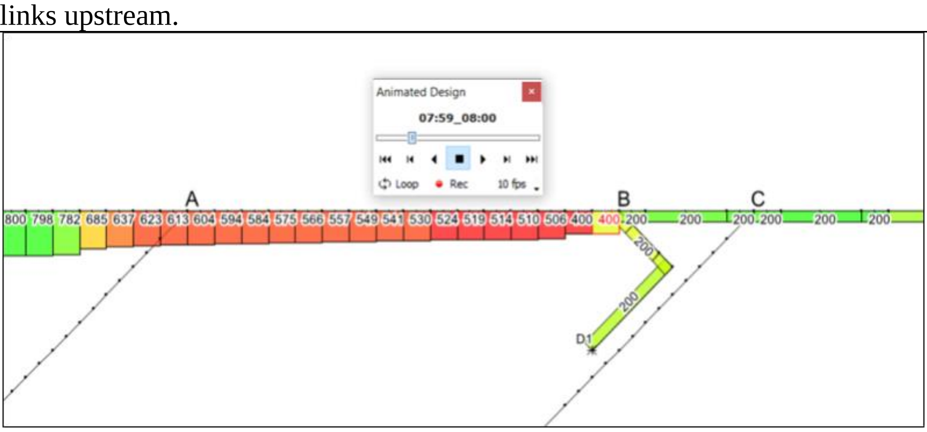
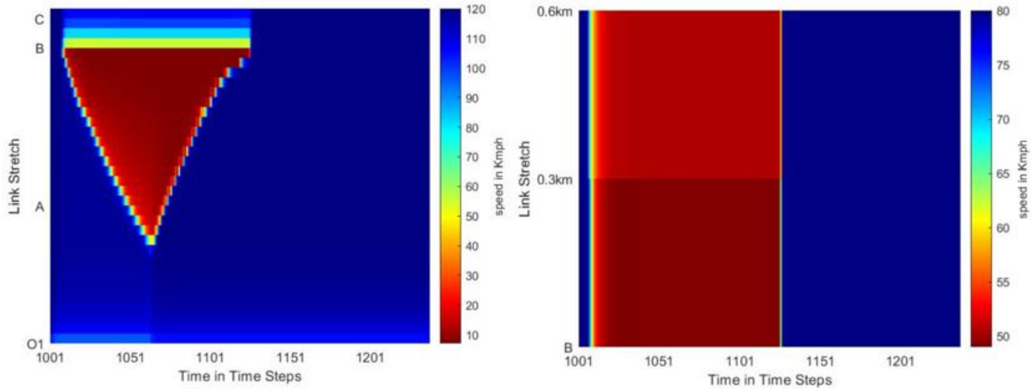
<table>
<tr><th>Test ID</th><th>2.3.1</th></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to check simple diverge model under free-flow conditions.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td>Node model-diverge behavior</td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1,</td></tr>
<tr><td>Demand Properties</td><td>
Table 22: Demand Matrices for Test No. 2.3.1
<table>
<tr><td></td><td>from</td><td>to</td><td></td><td>from</td><td>to</td><td></td><td>From</td><td>to</td><td></td><td>from</td><td>To</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>400</td><td>400</td><td>O1</td><td>400</td><td>400</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table>
</td></tr>
<tr><td>Expectation</td><td>Unrestricted flow is expected in all links of the highway and the off-ramp sections of the diverge links, as the links are undersaturated conditions.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>


Figure 96: Flow propagation chart for the highway corridor_2.3.1_StreamLine: MaDAM

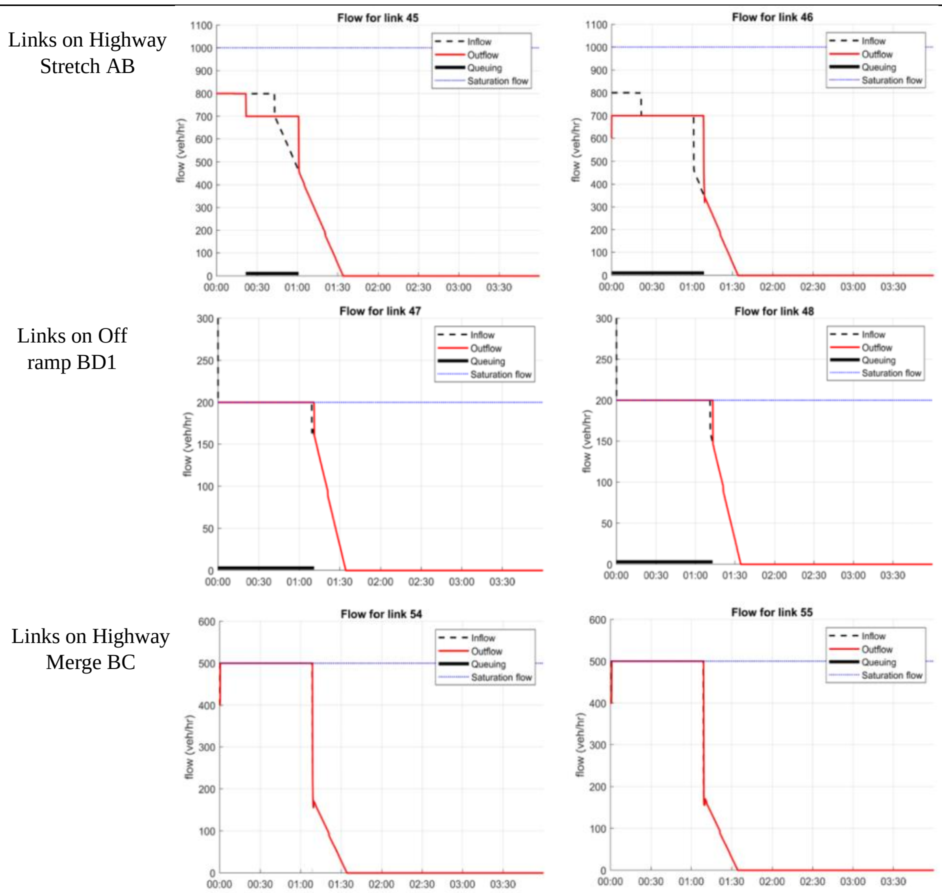
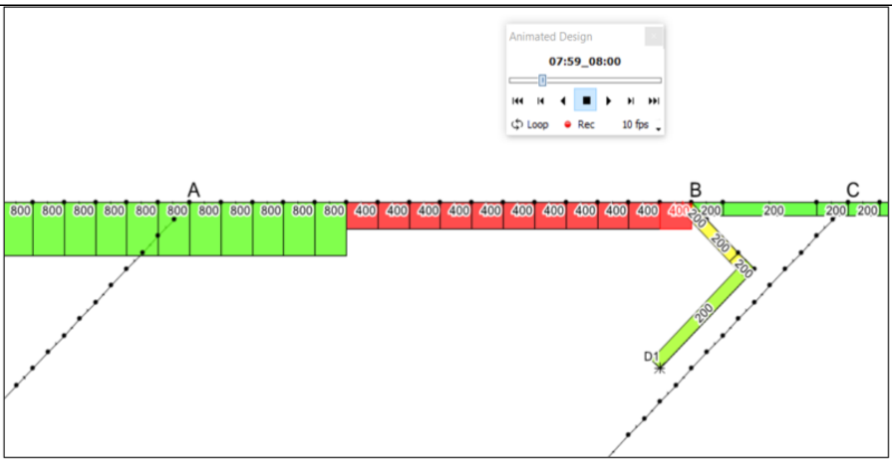
Expectation matches the simulation results. Exactly 400 veh/hr reaches both D1 and D2. This is shown through the propagation diagrams. Congestion is obsolete, as variables are free-flow state, throughout the simulation. In the last hour of the simulation, certain links showed abnormally low load values instead of a null value. We presume this could be error in simulation, during the last hour of the propagation, the network is practically empty.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
</table>

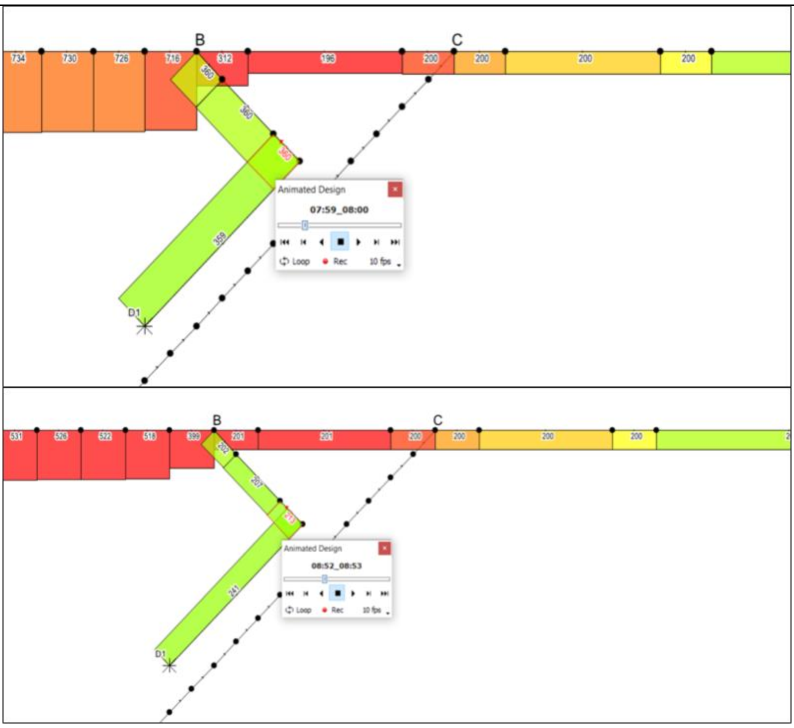
Results - MARPLE	 *Figure 97: Flow propagation values in links of the highway corridor_2.1.2_MARPLE* Propagation Behavior is in line with the expectation. Congestion is absent and the diverge nodes split the demand into the the diverging links to result in downstream flow values of 400 veh/hr to D1 and D2, as illustrated in Figure 97.
Score in EMMa	4
Results-StreamLine: eGLTM	 *Figure 98: Flow propagation chart for the highway corridor_2.3.1_StreamLine: MaDAM* Expectation matches the simulation results. Exactly 400 veh/hr reaches both D1 and D2. This is shown through the propagation diagrams. Congestion is obsolete, as variables are in free-flow state, throughout the simulation.
Score in EMMa	4
Test ID	**2.3.2**
Test Network	Same as 2.1.1

<table>
<tr><td>Test Description</td><td>The objective of this test is to check simple diverge model when the capacity of the receiving link is the constraint.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-diverge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1, the capacity of the links in off ramp stretch BD1 is reduced to 200 veh/hr</td></tr>
<tr><td>Demand Properties</td><td><i>Table 23: Demand Matrices for Test No. 2.3.2</i>
<table>
<tr><td></td><td>from</td><td>to</td><td></td><td>from</td><td>to</td><td></td><td>From</td><td>to</td><td></td><td>from</td><td>To</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>400</td><td>400</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>The capacity restriction in one of the diverge receiving link (off ramp links in BD1) is expected to reduce the outflow in the highway receiving link (in highway stretch BC), to the tune of outflow in BD1, to ensure FIFO. This outflow restriction, is further expected to cause congestion in the highway links upstream.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>


<i>Figure 99: Flow propagation chart for the highway corridor_2.3.2_StreamLine: MaDAM</i>


<i>Figure 100: Space Time Diagram of Highway Stretch (Left) and Off-ramp stretch BD1(Right)_2.3.2_StreamLine: MaDAM</i>
As expected, a congestion is found in the AB link as a result of capacity constraint in BD1 links. To obey the diverge node constraints, the outflow at B towards D1 is restricted to 200 veh/hr. This means 50% of vehicles arriving at B is restricted. Hence the total outflow at B will be 200/50% = 400 veh/hr. Which becomes the result of the simulation, as shown in the propagation diagram in Figure 99. The congestion pattern is also shown in the X-T diagram in Figure 100.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
</table>

Results - MARPLE	 *Figure 101: Flow propagation values in links of the highway corridor_2.3.2_MARPLE* Results are not in line with the expectation. As FIFO principle needs to be respected in diverge, we expect queue on AB, sending 200 to BD1 and 200 to BC. However, the results show an additional 300 veh/hr in the AB stretch to give a total outflow of 700 veh/hr which is not theoretically consistent (ref to Figure 101). The additional 300 veh/hr makes the outflow of the highway stretch BC to be 500 veh/hr which cannot be explained by the theoretical expectation of a diverge node, as demand itself is only 400 veh/hr.
Score in EMMa	1
Results-StreamLine: eGLTM	 *Figure 102: Flow propagation chart for the highway corridor_2.3.2_StreamLine: eGLTM* As expected, a congestion is found in the links in AB as a result of capacity constraint in BD. To obey the diverge node constraints, the outflow at B towards D1 is restricted to 200 veh/hr. This means 50% of vehicles arriving at B is restricted. Hence the reduced capacity of the upstream link

<table>
<tr><td></td><td>from B in AB is 200/50% = 400 veh/hr. The inflow of 400 + 400 vehicles causes the congestion build up in the stretch AB. The length of spillback is lesser compared to MaDAM due to lack of second order effects which is as per expectation (Ref to propagation chart in Figure 102).</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Test ID</td><td>2.3.3</td></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to check simple diverge model behavior in the event of a congestion and a spillback from a bottleneck downstream of the diverge node.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-diverge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1, the capacity of the links in highway stretch CD2 is reduced to 200 veh/hr</td></tr>
<tr><td>Demand Properties</td><td><i>Table 24: Demand Matrices for Test No. 2.3.3</i>
<table>
<tr><th></th><th>from</th><th>to</th><th></th><th>from</th><th>to</th><th></th><th>From</th><th>to</th><th></th><th>from</th><th>To</th></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>400</td><td>400</td><td>O1</td><td>400</td><td>400</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>The expectation here would be that the reduced capacity in the links in CD2 would trigger spillback upstream as a result of queuing and diverge node behavior such as FIFO is expected to follow at location B.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

<i>Figure 103: Flow propagation chart for the highway corridor_2.3.3_StreamLine: MaDAM</i>

Results consistent with expectation, but only during the second hour of the demand propagation. The spillback from the stretch CD2, makes the outflow in BC to be 200veh/hr. The diverge node at B, satisfies the nodal constraints by sending the same flow value to BD1 as observed in the propagation chart in Figure 103.</td></tr>
</table>

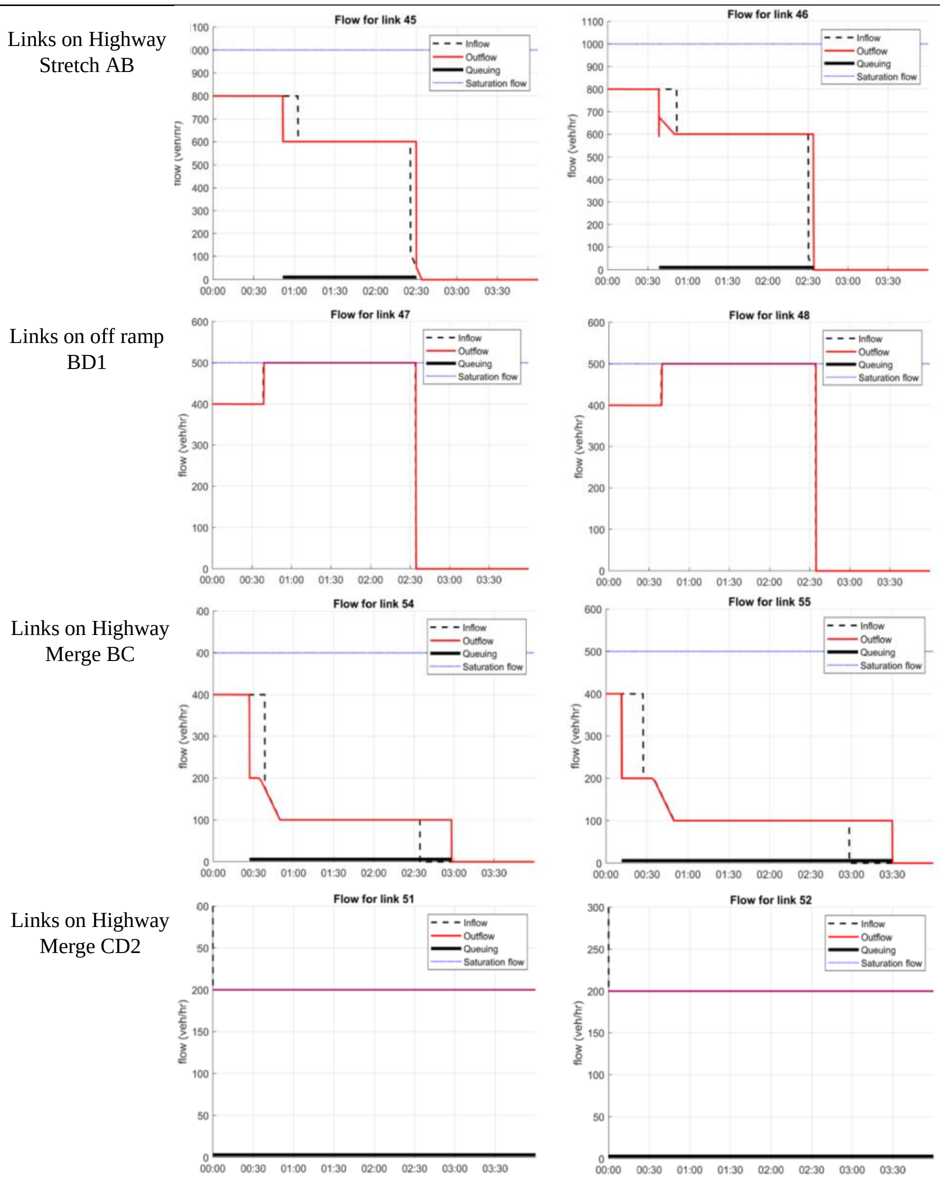
<table>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Results - MARPLE</td><td>


Figure 104: Flow propagation values in links of the highway corridor_2.3.3_MARPLE

We would expect here would be that the spillback from the link CD2 as a result of queuing would yield the links upstream to the same flow of 200 veh/hr in BC and BD2. For validating the diverge node behavior, the highway stretch in AB would be having an outflow of 400 veh/hr, which is split equally among the diverge links. However, the results are different from expectation. The congestion in BC as an effect of bottleneck in CD2, causes the speeds to drop drastically over the entire simulation period. This results in the link flows in this stretch becomes 100 veh/hr, which is not logical. At diverge node B, the outflow values in the off-ramp links in BD1 become 500 veh/hr. Rather than dropping to 200 veh/hr, the values strive to reach the capacity which is not consistent.
</td></tr>
<tr><td>Score in EMMa</td><td>1</td></tr>
</table>

<table>
<tr><td>Results-StreamLine: eGLTM</td><td>

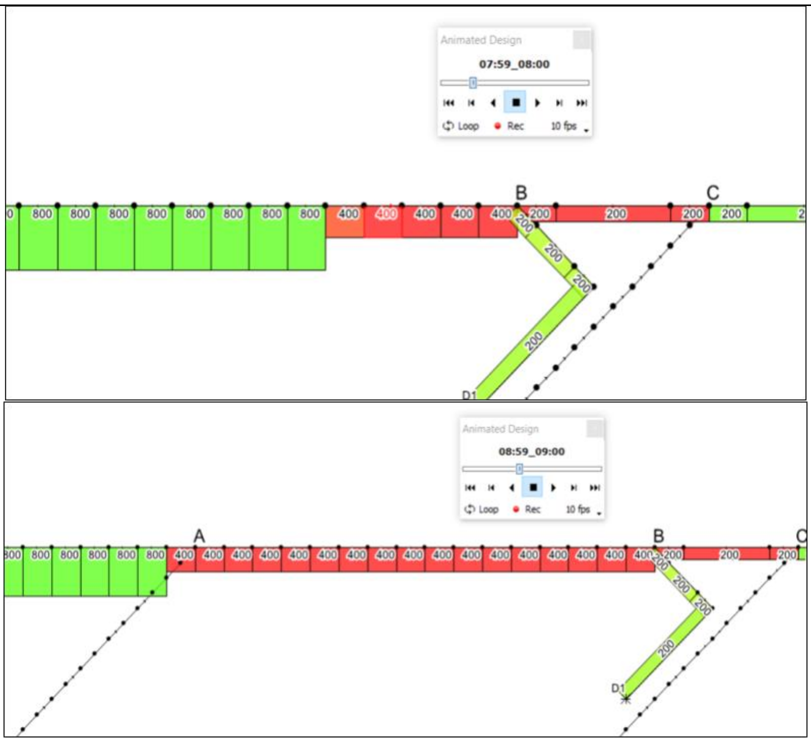


Figure 105: Flow propagation values in links of the highway corridor_2.3.3_eGLTM

Results consistent with expectation. The spillback from the stretch CD2, makes the outflow in BC to be 200veh/hr. The diverge node at B, satisfies the nodal constraints by sending the same outflow value to BD2 as observed in the propagation chart in Figure 105. The links in stretch AB has obtained reduced capacity which is the cause of the congestion. In contrast to the previous test case, the spillback is more evident in case of eGLTM compared to MaDAM, for the second hour of demand as it grows all the way to node A and beyond, completely in line with the expectation.</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Test ID</td><td>2.4.1</td></tr>
<tr><td>Test Network</td><td>Same as 2.1.1</td></tr>
<tr><td>Test Description</td><td>To test the consistency of turning rates at the diverge are always consistent with the O-D table, even when (time-dependent) delays occur in the network.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Node model-diverge behavior</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 2.1.1, capacity of links in off-ramp stretch BD1 is reduced to 400 veh/hr.</td></tr>
<tr><td>Demand Properties</td><td>Table 25: Demand Matrices for Test No. 2.4.1
<table>
<tr><td></td><td>from</td><td>to</td><td></td><td>from</td><td>to</td><td></td><td>From</td><td>to</td><td></td><td>from</td><td>To</td></tr>
<tr><td></td><td>07:00</td><td>08:00</td><td></td><td>08:00</td><td>09:00</td><td></td><td>09:00</td><td>10:00</td><td></td><td>09:00</td><td>10:00</td></tr>
<tr><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td><td></td><td>D1</td><td>D2</td></tr>
<tr><td>O1</td><td>100</td><td>300</td><td>O1</td><td>300</td><td>150</td><td>O1</td><td>0</td><td>0</td><td>O1</td><td>0</td><td>0</td></tr>
<tr><td>O2</td><td>300</td><td>100</td><td>O2</td><td>100</td><td>250</td><td>O2</td><td>0</td><td>0</td><td>O2</td><td>0</td><td>0</td></tr>
<tr><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td><td>O3</td><td>0</td><td>0</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>Congestion is expected in the network as the links in CD2 are over saturated during the demand simulation. The diverge node B will reduce the outflow in BD1 to the reduced capacity of links in BC as a consequence of spillback and ensuring FIFO. These diverge node constraints is expected to ensure that the complete demand propagated from O1 and O2, reaches D1 and D2, within the simulation period.</td></tr>
</table>

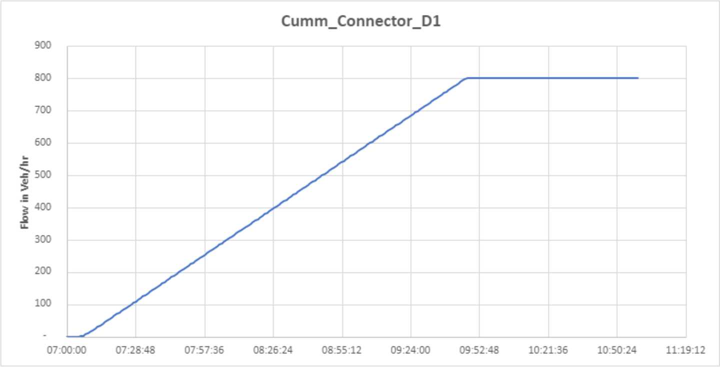
<table>
<tr>
<td>Results - StreamLine: MaDAM</td>
<td>


*Figure 106: Cumulative flow propagation values in connector links to destination **D1**_2.4.1_MaDAM*

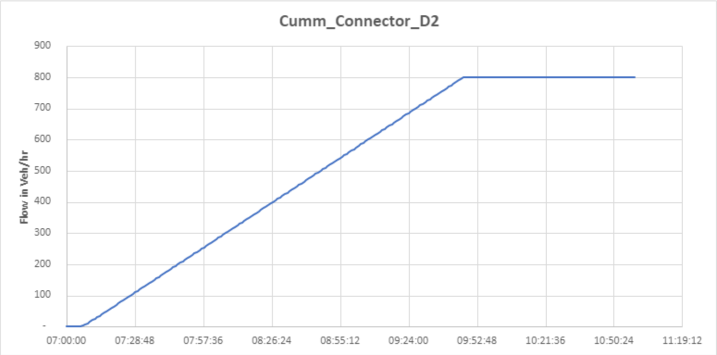


*Figure 107: Cumulative flow propagation values in connector links to destination **D2**_2.4.1_MaDAM*

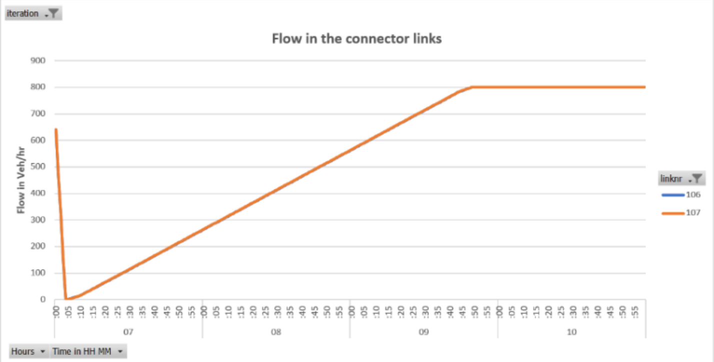
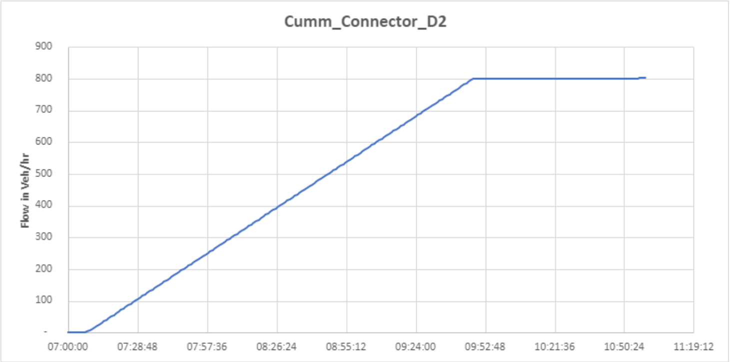
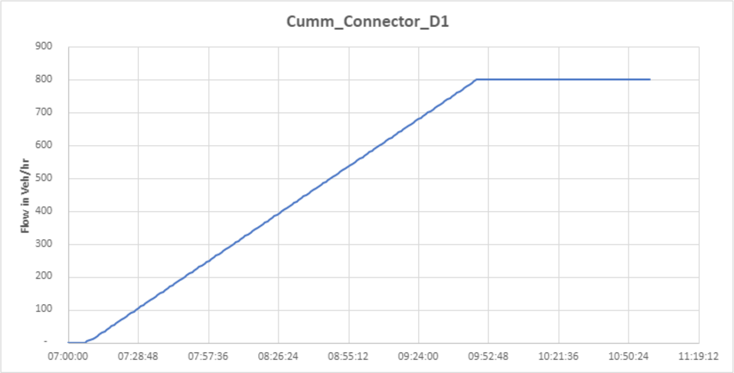
The expectation is that the both the destinations receive the vehicles as per the OD matrices. Which in this case would be 400 veh/hr X 2 hrs of demand = 800 veh/hr. This is indeed the case confirmed with the cumulative plots in Figure 106 and Figure 107.
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>4</td>
</tr>
<tr>
<td>Results - MARPLE</td>
<td>


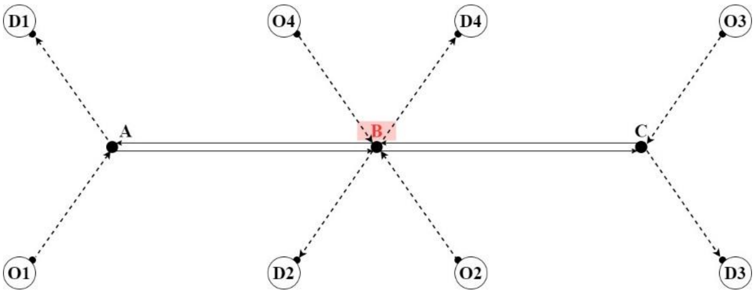
*Figure 108: Cumulative flow propagation values in connector links to destination **D1(106)** and **D2(107)**_2.4.1_MaDAM*

Results match expectation, as shown in Figure 108
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>4</td>
</tr>
</table>

Results-StreamLine: eGLTM	 *Figure 109: Cumulative flow propagation values in connector links to destination **D2**_2.4.1_eGLTM*  *Figure 110: Cumulative flow propagation values in connector links to destination **D1**_2.4.1_eGLTM* Results match expectation, as shown in Figure 109 and Figure 110.Figure 108
Score in EMMa	**4**

3. Signalized intersection– urban network

The series of tests are conducted to understand the behavior of the model in an urban intersection(signalized). The test series can also be extended to include other types of controlled intersections such as roundabouts. The test networks mainly adopted from (FakhraeiRoudsari, Huang, & Tampère, 2015).

Test ID	**3.1.1**
Test Network	 No of Links Between AB – 2 (Link Nr 1 and Link Nr 2) No of Links Between BC – 2 (Link Nr 3 and Link Nr 4)
Test Description	The objective of this test is to understand the intersection behavior when the links are in under-saturated conditions
MoP Evaluated -	*Signalized Intersection*

<table>
<tr><td>Link to EMMa</td><td></td></tr>
<tr><td>Supply Properties</td><td>
Table 26: Network properties for Test No. 3.1.1
<table>
<tr><th></th><th>Capacity</th><th>Length (km)</th><th>Free Flow Speed</th><th>Speed at Cap</th><th>No. of Lane</th></tr>
<tr><td>All Links</td><td>1000</td><td>1</td><td>60</td><td>35</td><td>1</td></tr>
<tr><td>Connectors</td><td>1000</td><td>1</td><td>60</td><td>35</td><td>1</td></tr>
</table>
Saturation flow of the intersection B = 1000 Veh/hr

<u>Signal Timings</u>
Cycle time: 120 Seconds
Green time: 60 Seconds
</td></tr>
<tr><td>Demand Properties</td><td>
Table 27: Demand Profile for Test No. 3.1.1
<table>
<tr><th>Sl No</th><th>Time</th><th>Demand (O1-D3)</th><th>Saturation %</th></tr>
<tr><td>1</td><td>7-8</td><td>125</td><td>25%</td></tr>
<tr><td>2</td><td>8-9</td><td>350</td><td>70%</td></tr>
<tr><td>3</td><td>9-10</td><td>425</td><td>85%</td></tr>
<tr><td>4</td><td>10-11</td><td>490</td><td>98%</td></tr>
<tr><td>5</td><td>11-12</td><td>0</td><td>Empty</td></tr>
</table>
</td></tr>
<tr><td>Expectation</td><td>Unrestricted movement of traffic is expected in the links between A and B, as links are in undersaturated conditions throughout the simulation.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td></td></tr>
</table>

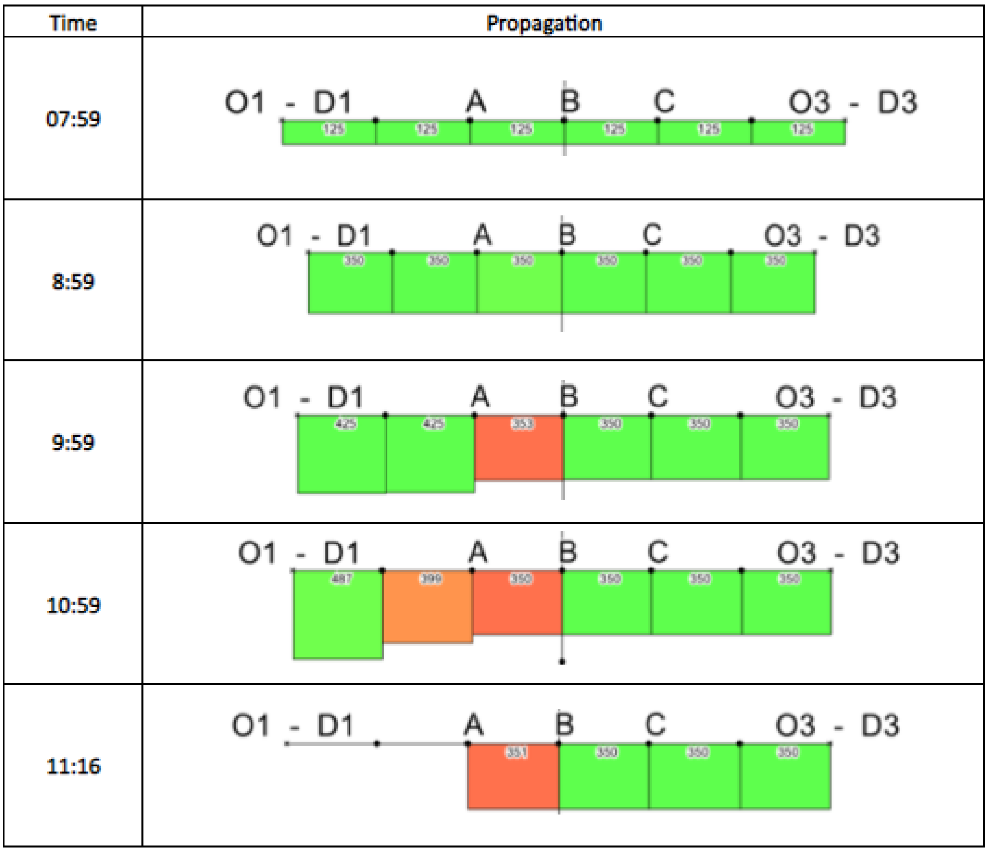


Figure 111: Flow propagation values in links of the urban intersection network_3.1.1_MaDAM

Figure 112: Route travel cost in the urban road stretch AB throughout the simulation_3.1.1_MaDAM

The expectation was that the turning and flow propagation would take place without congestion. However, the results do not seem to match the expectation. As shown in the propagation diagram in Figure 111, a congestion is developed at the 9-10 demand duration when the saturation is 85%. The junction model in StreamLine reduces the link capacity by an arbitrary 300 veh/hr. This would mean the new capacity is 700 veh/hr. As the green time is 50% of the cycle time, the reduced outflow at the turn would be 350 veh/hr. As result of this capacity reduction, it makes sense that rest of the downstream links propagate traffic at a flow rate of 350 veh/hr till the duration of the demand input. On the basis of this explanation, the trend observed for travel costs over the links is plotted in the line chart shown in Figure 112. It can be observed that the travel cost (time) is increasing or the delay is increasing even in

<table>
<tr><td></td><td>the undersaturated condition, when the saturation increases.</td></tr>
<tr><td>Score in EMMa</td><td>**2**</td></tr>
<tr><td>Results - MARPLE</td><td>

Table 28: Route travel cost in the urban road stretch AB throughout the simulation_3.1.1_MARPLE

<table>
<tr><th>Time period</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th></tr>
<tr><td>**Route Cost**</td><td>6.32</td><td>6.47</td><td>6.58</td><td>6.61</td><td>6.21</td></tr>
</table>

Links Upstream of Intersection B, in Stretch AB

Links Downstream of Intersection B, in Stretch BC

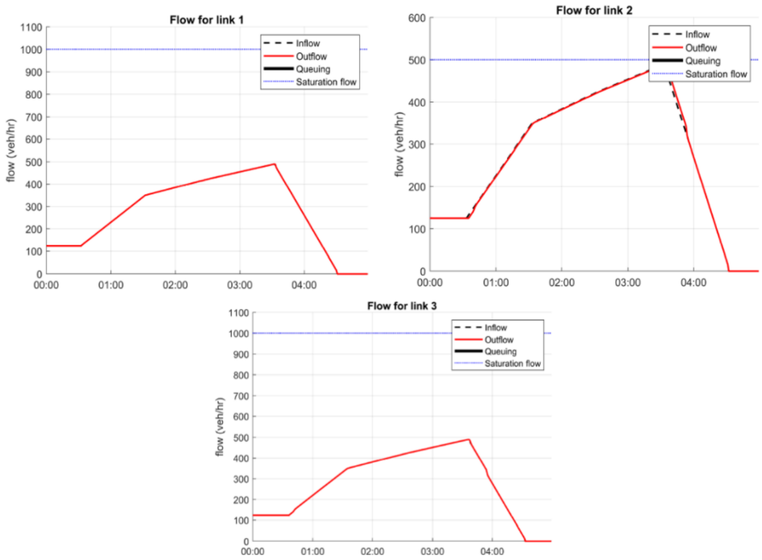


Figure 113: Flow propagation values in links of the urban intersection network_3.1.1_MARPLE

The new capacity as an effect of the signal control is 50%*saturation flow * no of lanes, which becomes 500 veh/hr. This test loads only one turn without conflicting traffic until capacity. We expect increasing travel time (delay) when demand is increasing from 25% saturation till 98% (especially in near-saturation, i.e. from 80% and up). No congestion should occur. This is indeed the case as observed through the flow charts (Figure 113) and route travel costs (Table 28).

</td></tr>
<tr><td>Score in EMMa</td><td>**4**</td></tr>
</table>

<table>
<tr>
<td>Results-StreamLine: eGLTM</td>
<td>

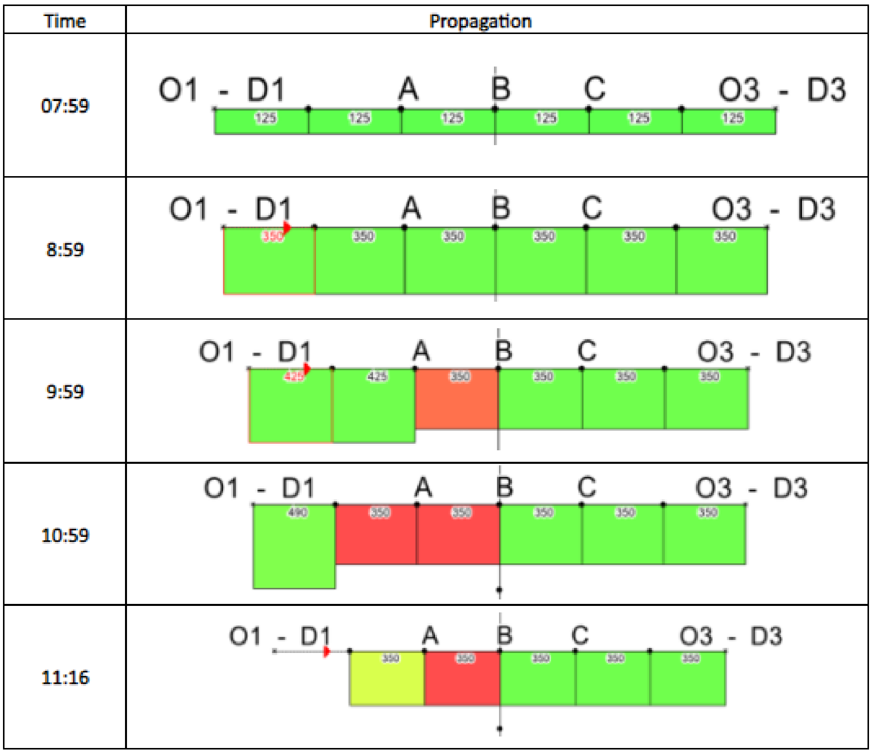


Figure 114: Flow propagation values in links of the urban intersection network_3.1.1_eGLTM

Figure 115: Travel cost in the urban road stretch AB throughout the simulation_3.1.1_eGLTM

Results are similar to that of MaDAM. This is expected as both MaDAM and eGLTM are built in the StreamLine framework.
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>2</td>
</tr>
<tr>
<td>Test ID</td>
<td>3.1.2</td>
</tr>
<tr>
<td>Test Network</td>
<td>Same as 3.1.1</td>
</tr>
<tr>
<td>Test Description</td>
<td>The objective of this test is to understand the intersection behavior when the links are in oversaturated conditions.</td>
</tr>
<tr>
<td>MoP Evaluated - Link to EMMa</td>
<td><i>Signalized Intersection</i></td>
</tr>
</table>

<table>
<tr><td>Supply Properties</td><td>Same as 3.1.1</td></tr>
<tr><td>Demand Properties</td><td>

Table 29: Demand Profile for Test No. 3.1.2

<table>
<tr><th>Sl No</th><th>Time</th><th>Demand (O1-D3)</th><th>Saturation %</th></tr>
<tr><td>1</td><td>7-8</td><td>600</td><td>120%</td></tr>
<tr><td>2</td><td>8-9</td><td>0</td><td>Empty</td></tr>
<tr><td>3</td><td>9-10</td><td>0</td><td>Empty</td></tr>
<tr><td>4</td><td>10-11</td><td>0</td><td>Empty</td></tr>
<tr><td>5</td><td>11-12</td><td>0</td><td>Empty</td></tr>
</table>

</td></tr>
<tr><td>Expectation</td><td>The turn flow restriction at the intersection is expected to reduce the capacity by 50% as green time/cycle time is 50%. This in turn is expected to reduce the outflow of link AB to 500 veh/hr and thereby create congestion and subsequent spillback in the route between O1 and D3.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

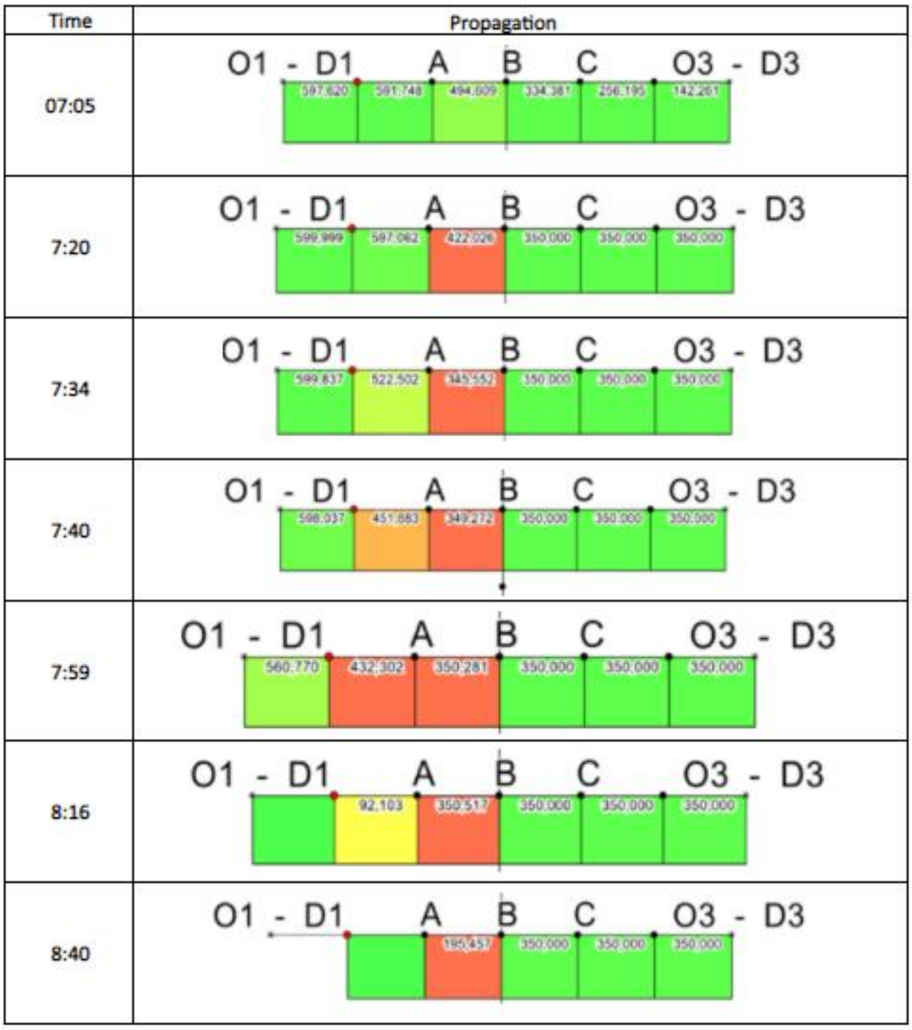


Figure 116: Flow propagation values in links of the urban intersection network_3.1.2_MaDAM

</td></tr>
</table>

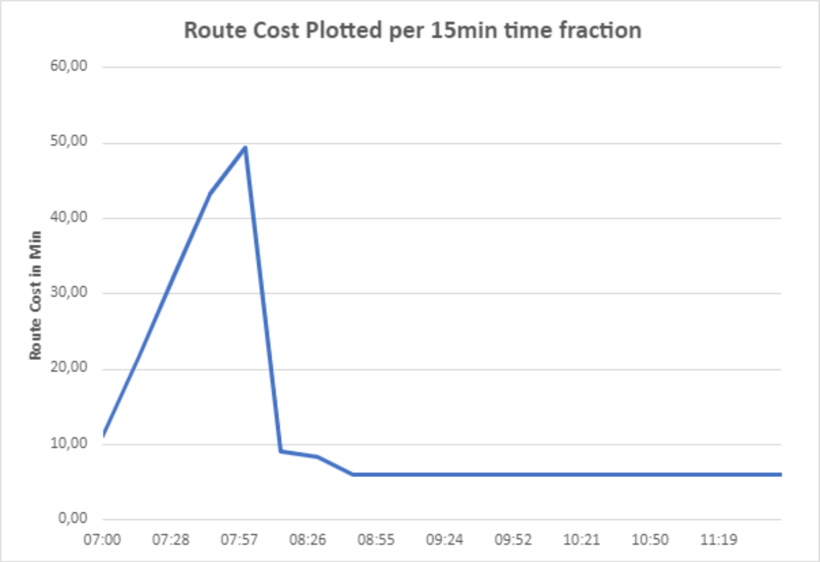


Figure 117: Travel cost in the urban road stretch AB per route choice time fraction of 15mins interval_3.1.2_MaDAM

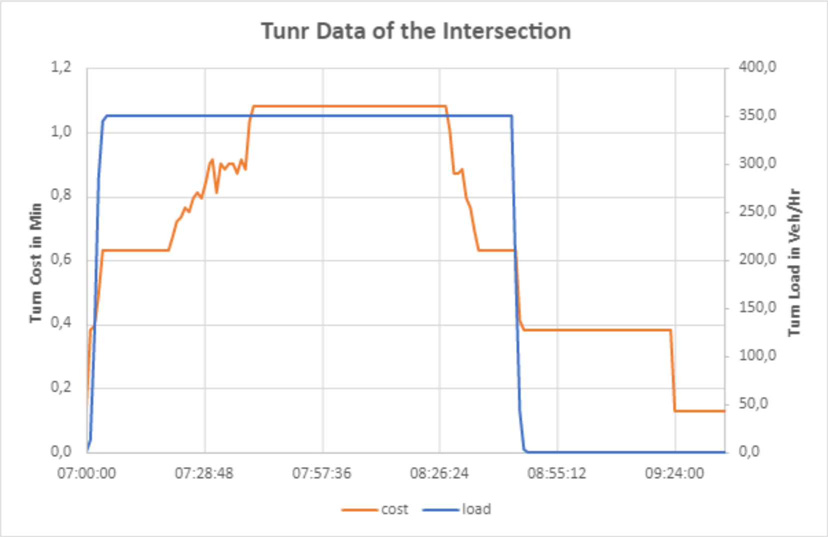


Figure 118: Turn data of the urban intersection B, throughout the network loading phase_3.1.2_MaDAM

The expected congestion is less than the simulated results. As a result of the junction model XStream, there is a drop in the capacity by 300 veh/hr. Hence the outflow capacity of the intersection remains at 350 veh/hr as explained in the previous test. Due to oversaturated conditions, the congestion occurs within the propagation duration and spills back to the links upstream and thereby result in a speed reduction. The capacity reduction has caused a more severe congestion as shown in Figure 116. There were some anomalous error values (abnormally low values) in the load and densities during emptying the network, which were identified and removed from the results. StreamLine: MaDAM stores the turn data of the intersection as a separate object, which provides the turn cost (travel time in minutes required to traverse the intersection) and turn capacity/load values (maximum flow value allowed to traverse the intersection) throughout the simulation as shown in Figure 118. In general, the values are as per the expectation, where the turn cost values and the travel time values increases as the congestion builds up and total delay increases. However, we observe a series of cost fluctuation from 07:20, which is also abnormal. Looking at the route cost graph, it is as expected that the route cost is at its peak when the congestion is peak at 07:59. The route travel time reduction occurs, when the flow is less than capacity, at around 08:40.

Score in EMMa	2

Results - MARPLE

Table 30: Route travel cost in the urban road stretch AB throughout the simulation_3.1.2_MARPLE

Time Period	**1**	**2**	**3**	**4**	**5**
Route Cost	6.32	6.47	6.58	6.61	6.21

Links Upstream of Intersection B, in Stretch AB

Links Downstream of Intersection B, in Stretch BC

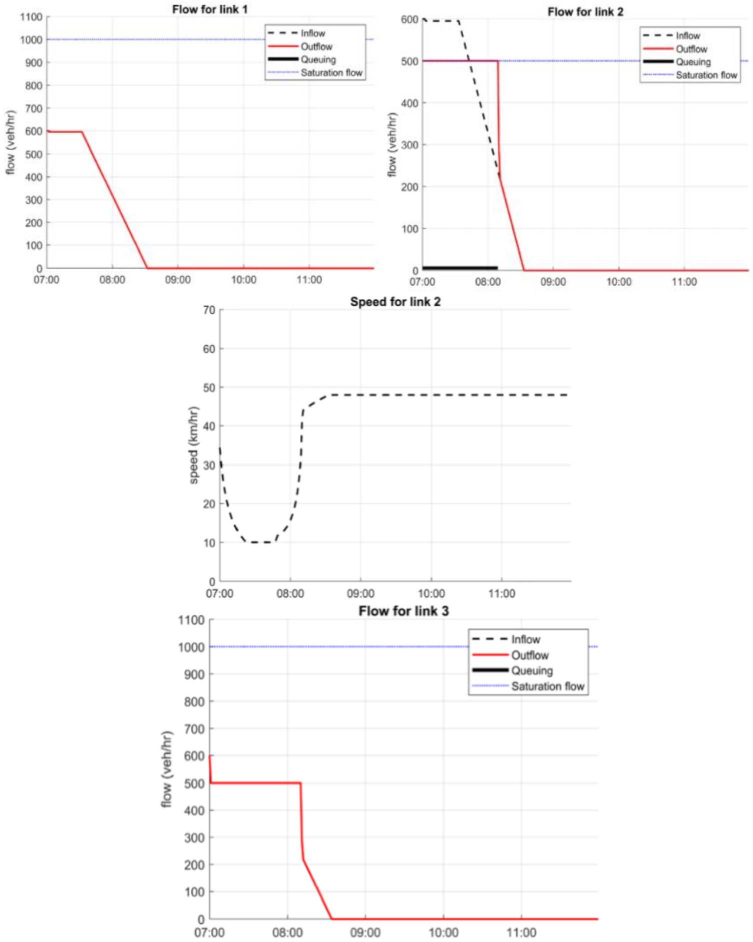


Figure 119: Flow values and speed values in links of the urban intersection network_3.1.2_MARPLE

The expected congestion in the controlled link connected to the intersection is high, with speed dropping to a low value, indicating an almost standstill in the oversaturated condition as shown in Figure 119. However, the issues encountered during the propagation tests with regard to spillback is present in this test. The links upstream of the congested link does not encounter any speed drop or queuing, which implies the absence of spillback. The trend for route travel costs per iteration is as per expectation, as shown in Table 30.

Score in EMMa

3

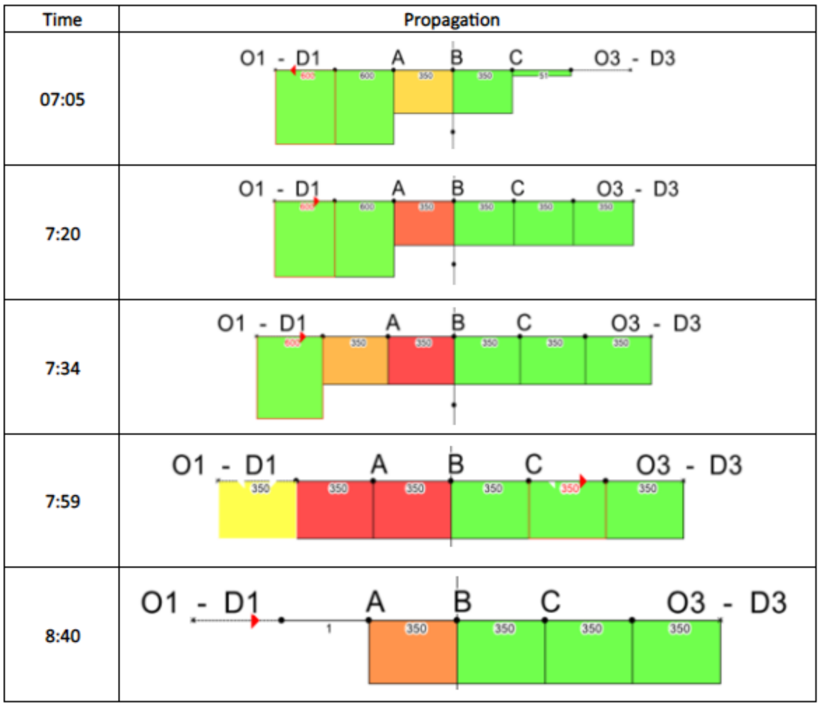
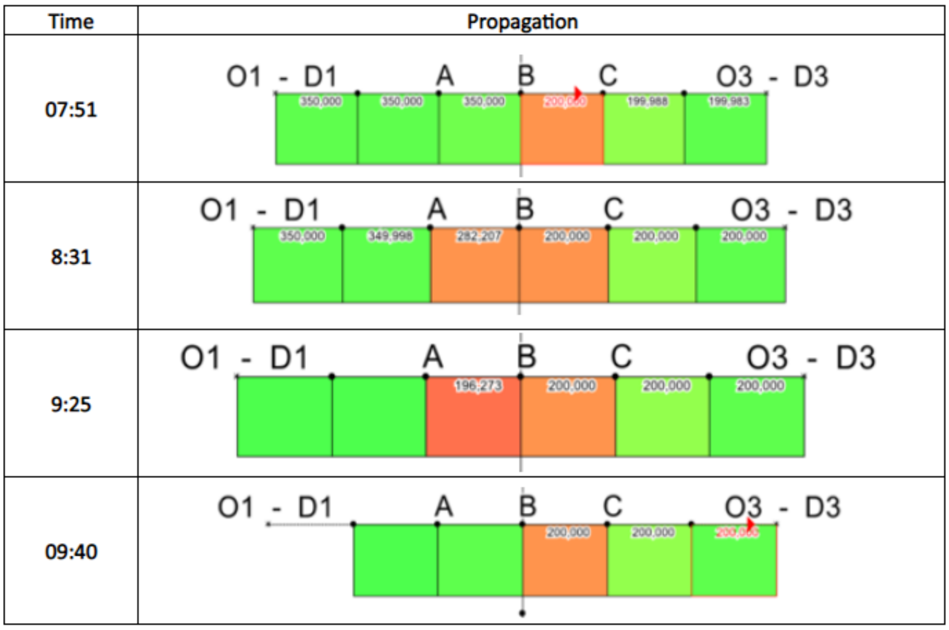
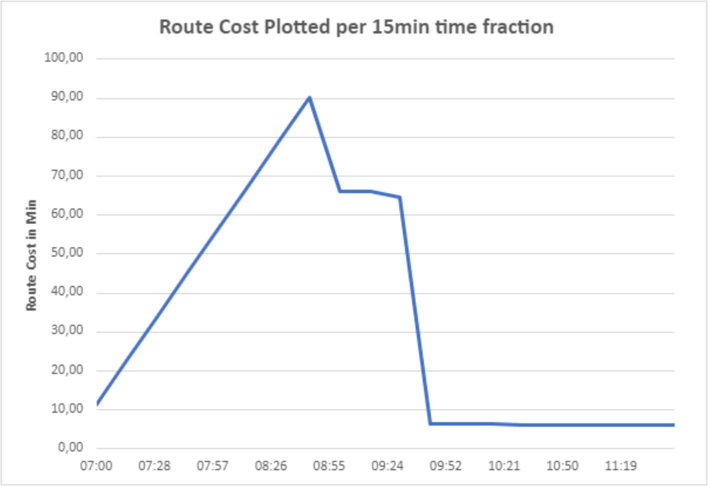
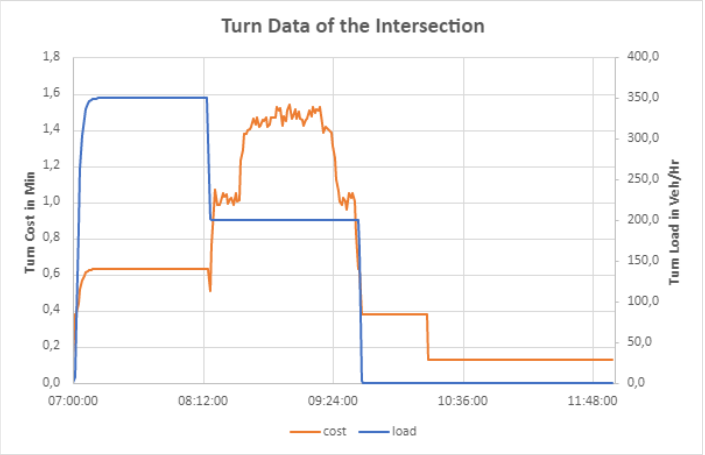
<table>
<tr><td>Results-StreamLine: eGLTM</td><td>


Figure 120: Flow propagation values in links of the urban intersection network_3.1.2_eGLTM

Results are similar to that of MaDAM. It may be noted that there was an issue encountered in the turn data in StreamLine: eGLTM. The turn data value show loads of 1000 veh/hr as opposed to 350 veh/hr (based on the results from the previous tests), which is abnormal. The turn cost values are shown as nil as well. Interestingly the link propagation values are not affected by this and link loads are as per expectation from previous tests as shown in Figure 120. This may be inferred as an issue encountered in the storage and writing of turn data into the turn objects.
</td></tr>
<tr><td>Score in EMMa</td><td>2</td></tr>
<tr><td>Test ID</td><td>3.1.3</td></tr>
<tr><td>Test Network</td><td>Same as 3.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to understand the intersection behavior when the links are in oversaturated conditions due to congestion and consequent spillback from a bottleneck downstream.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td>Signalized Intersection</td></tr>
<tr><td>Supply Properties</td><td>Same as 3.1.1, except the capacity of link in stretch CO3 is reduced to 200 veh/hr</td></tr>
<tr><td>Demand Properties</td><td>
Table 31: Demand Profile for Test No. 3.1.3
<table>
<tr><th>Sl No</th><th>Time</th><th>Demand (O1-D3)</th><th>Saturation %</th></tr>
<tr><td>1</td><td>7-8</td><td>350</td><td>70%</td></tr>
<tr><td>2</td><td>8-9</td><td>350</td><td>70%</td></tr>
<tr><td>3</td><td>9-10</td><td>0</td><td>Empty</td></tr>
<tr><td>4</td><td>10-11</td><td>0</td><td>Empty</td></tr>
<tr><td>5</td><td>11-12</td><td>0</td><td>Empty</td></tr>
</table>
</td></tr>
</table>

Expectation	The congestion from the bottleneck downstream is expected to increase the delay in the route and also the turn cost of the intersection.
Results - StreamLine: MaDAM	 *Figure 121: Flow propagation values in links of the urban intersection network_3.1.3_MaDAM*  *Figure 122: Travel cost in the urban road stretch AB per route choice time fraction of 15mins interval_3.1.3_MaDAM*  *Figure 123: Turn data of the urban intersection B, throughout the network loading phase_3.1.3_MaDAM*

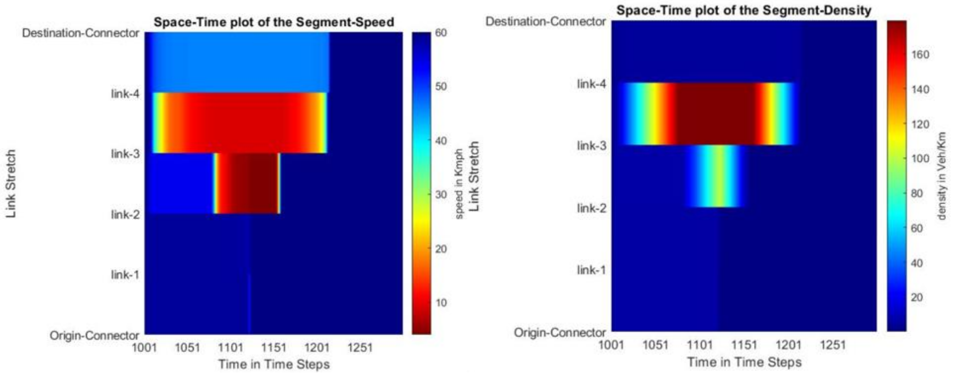


Figure 124: Space Time Diagram-Values based on Speed (Left) and Values based on Density (Right) _3.1.3_StreamLine: MaDAM

As expected, the congestion from the link downstream of the intersection affects the turn cost and turn flow as showcased in the turn data chart (Figure 122 and Figure 123). The turn cost is higher than the previous test case as a result of delays due to the bottleneck and subsequent spillback to the links upstream. As per expectation the turn cost values increase when the spillback occurs beyond the intersection after 08:00 AM as shown in Figure 123. At the same time, the turn flow is restricted to the capacity of the bottleneck, till the time the demand drops to zero and the network recovers from congestion at around 09:45. The spillback and congestion pattern is shown in the X-T diagrams of speed and density (Figure 124). The intersection acts as a normal node for passing on congestion upstream. Unusually, the density during congestion reaches almost the jam density of 180 veh/km. This is different from the test no 3.1.1, where the density values even in oversaturated conditions, never reached the jam density value.

Score in EMMa

3

Results - MARPLE

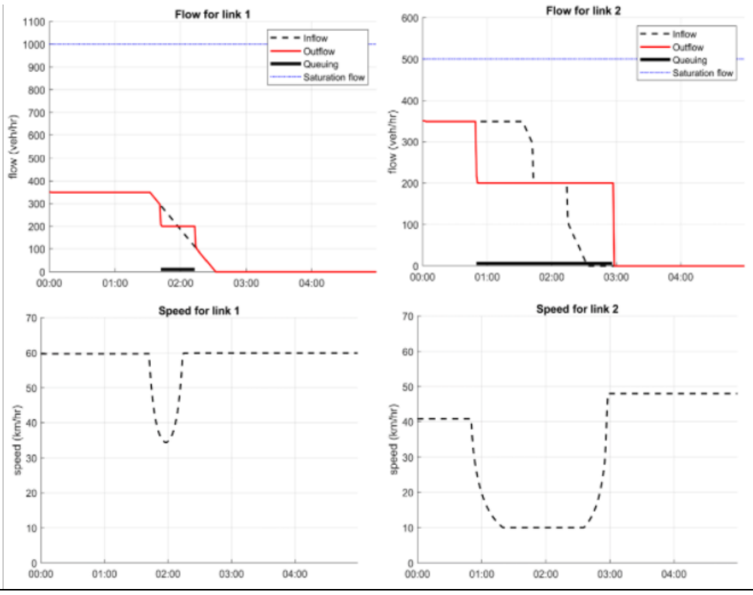
Table 32: Route travel cost in the urban road stretch AB throughout the simulation_3.1.3_MARPLE

Time period	**1**	**2**	**3**	**4**	**5**
Route Cost in min	11.44	16.86	16.45	9.04	6.25

Table 33: Route delay in the urban road stretch AB throughout the simulation_3.1.3_MARPLE

Time period	**1**	**2**	**3**	**4**	**5**
Route Delay in min	5.44	10.86	10.45	3.04	0.25

Links Upstream of Intersection B, in Stretch AB



<table>
<tr>
<td></td>
<td>
Links Downstream
of Intersection B,
in Stretch BC

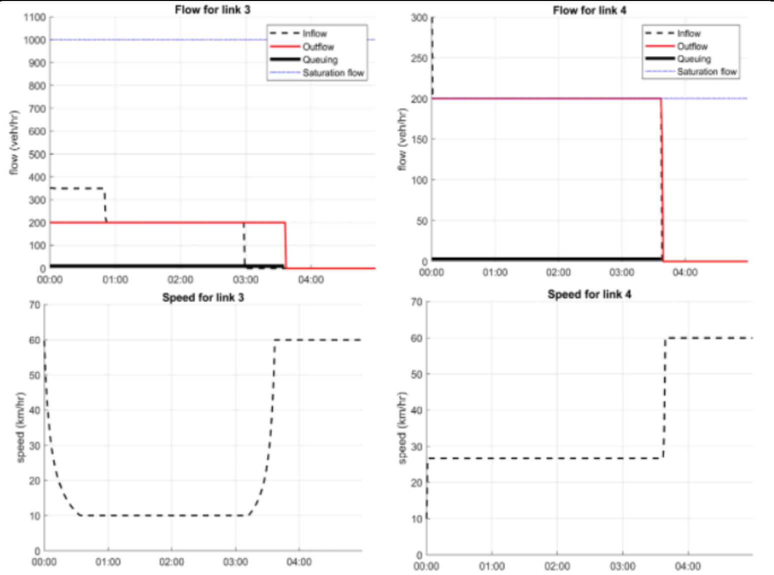


Figure 125: Flow values and speed values in links of the urban intersection network_3.1.3_MARPLE

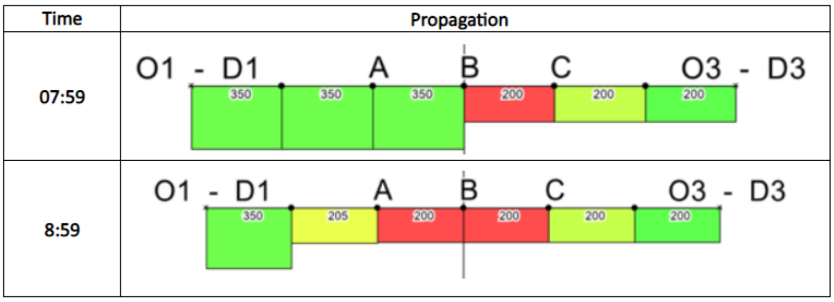
The intersection model captures the effect of the spillback from bottleneck downstream to produce delay in the controlled link nr 2. Furthermore, the effect of spillback trickles down to even link no-1. This means, the spillback effect is effectively captured by the model when the demand duration is sufficient. However, the queuing starts to occur in the bottleneck link nr 4, as opposed to link nr 3, which is theoretically not consistent (also different from the propagation tests in series 1). The queuing is also evident in the links upstream of the bottleneck due to spillback as shown in Figure 125. The route travel times and the route delays are also representative of influence of downstream bottleneck with increasing route travel times and delays in time period 2 and 3 compared to time period 1, as shown in Table 32 and Table 33.
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>3</td>
</tr>
<tr>
<td>Results-StreamLine: eGLTM</td>
<td>

<table>
<tr><th>Time</th><th>Propagation</th></tr>
<tr><td>07:59</td><td>O1 - D1   A   B   C   O3 - D3
350 350 350 200 200 200</td></tr>
<tr><td>8:59</td><td>O1 - D1   A   B   C   O3 - D3
350 205 200 200 200 200</td></tr>
</table>

Figure 126: Flow propagation values in links of the urban intersection network_3.1.3_eGLTM

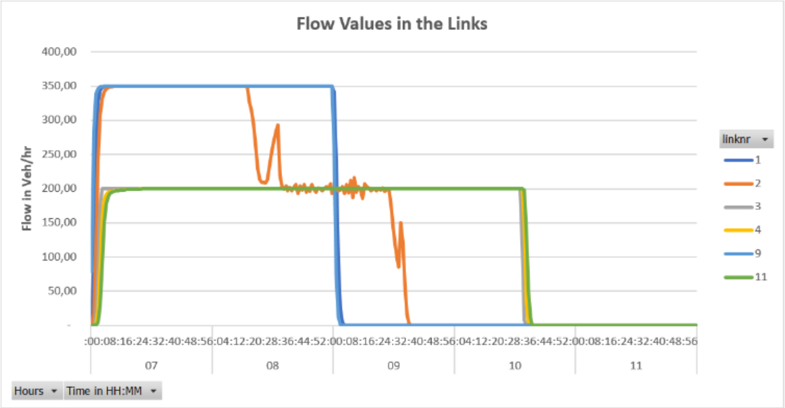
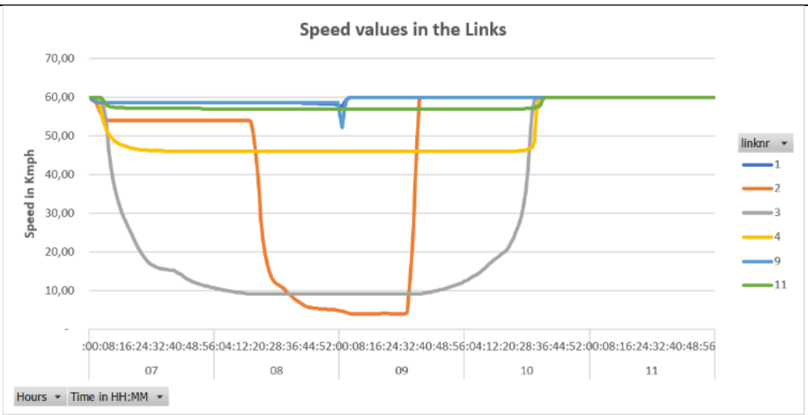
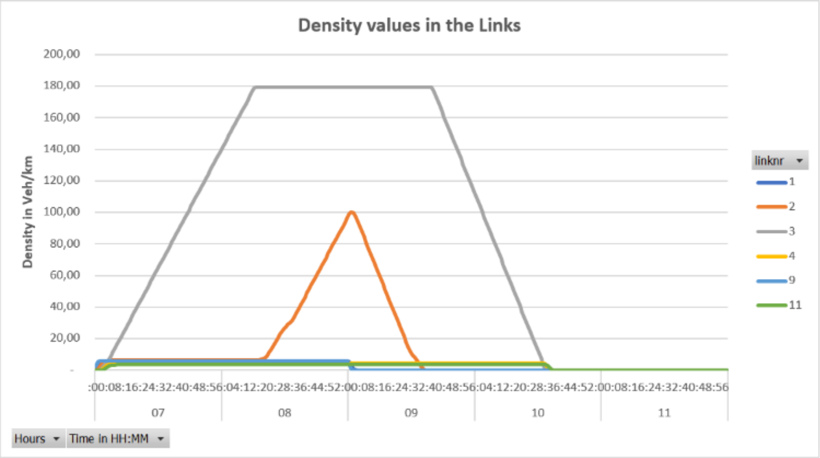


Figure 127: Flow values on the urban network_3.1.3_ eGLTM
</td>
</tr>
</table>

	 *Figure 128: Speed values on the urban network_3.1.3_ eGLTM*  *Figure 129: Density values on the urban network_3.1.3_ eGLTM* Results are somewhat in line with the expectation. However, some abnormalities are noticed in the propagation which shows inconsistent behavior. As shown in the flow propagation values in Figure 127, for link 2, before dropping down to a capacity 200 veh/hr as an effect of spill back from link 3, a peak is observed at around 08:25. Furthermore, from the speed chart in Figure 128 and the density chart in Figure 129, we notice that the vehicles come to a near standstill case in link 3, with density touching maximum value and speed dropping to almost 10km. Same is the case with link 2. Interestingly the speed drop here is more than that of link 3, to a value close to zero.
Score in EMMa	**2**
Test ID	**3.2.1**
Test Network	Same as 3.1.1
Test Description	The objective of this test is to understand the influence of the intersection in the route travel time when there is through traffic and diverging traffic at the intersection (undersaturated condition).
MoP Evaluated - Link to EMMa	*Signalized Intersection*
Supply Properties	Same as 3.1.1, the turn saturation of the intersection is increased to 1300 veh/hr(based on results of previous three tests).

<table>
<tr><td>Demand Properties</td><td>

Table 34: Demand Profile for Test No. 3.2.1

<table>
<tr><th>Time</th><th>Demand (O1-D3)</th><th>Demand (O1-D2)</th><th>Demand (O1-D4)</th><th>Sum</th><th>Saturation %</th><th>Expectation</th></tr>
<tr><td>7-8</td><td>95</td><td>15</td><td>15</td><td>125</td><td>25%</td><td>Small Delay</td></tr>
<tr><td>8-9</td><td>300</td><td>25</td><td>25</td><td>350</td><td>70%</td><td>Small Delay</td></tr>
<tr><td>9-10</td><td>300</td><td>25</td><td>100</td><td>425</td><td>85%</td><td>Delay</td></tr>
<tr><td>10-11</td><td>400</td><td>45</td><td>45</td><td>490</td><td>98%</td><td>Heavy Delay</td></tr>
<tr><td>11-12</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
</table>

</td></tr>
<tr><td>Expectation</td><td>We would expect unrestricted propagation of traffic throughout the simulation, with a gradual increase in route travel cost over time due to increasing saturation percentages.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

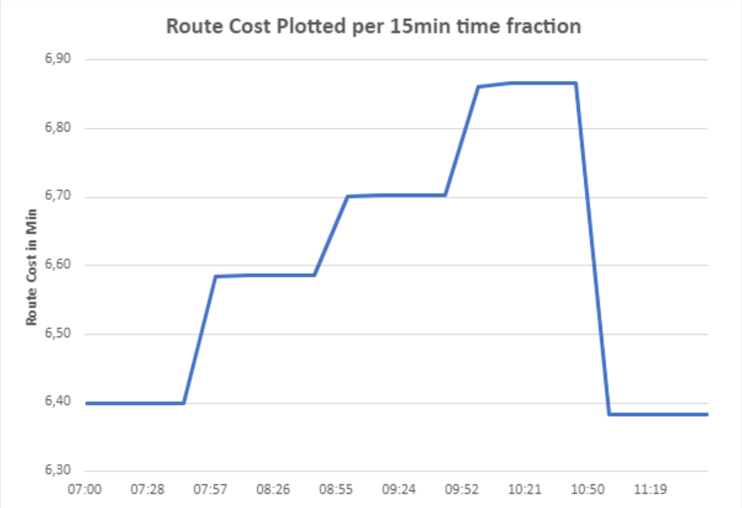


Figure 130: Travel cost in the urban road stretch AB per route choice time fraction of 15mins interval_3.2.1_MaDAM

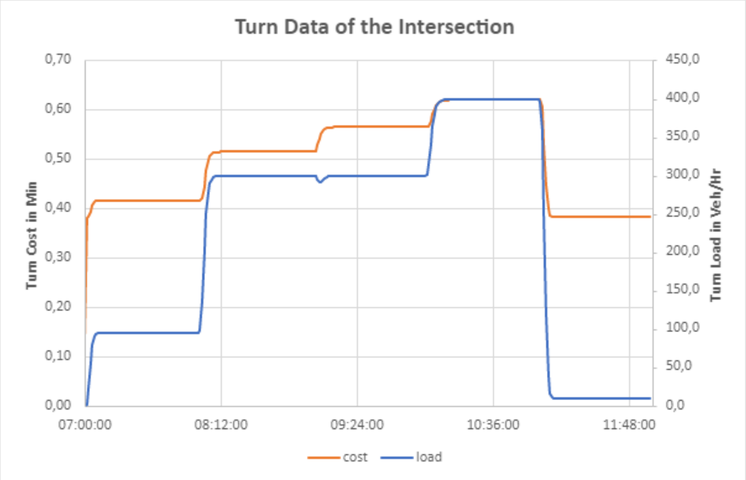
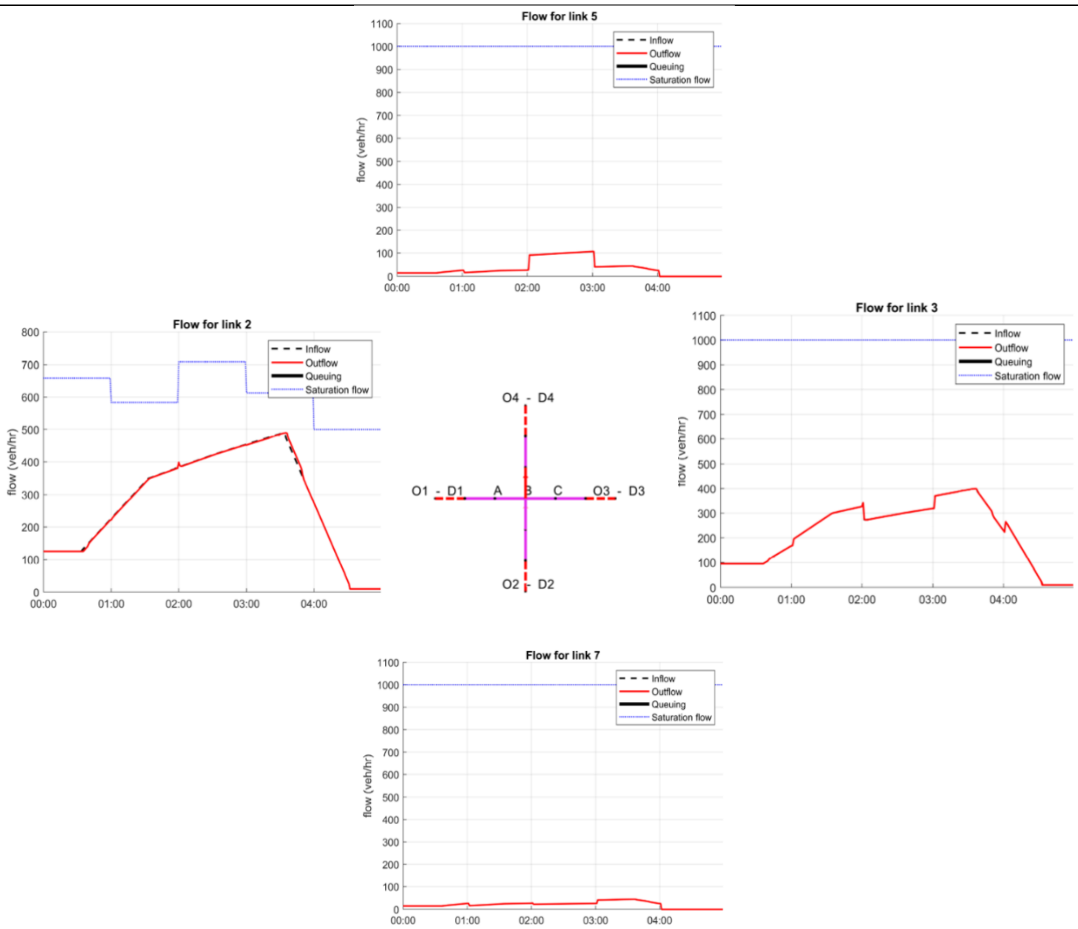
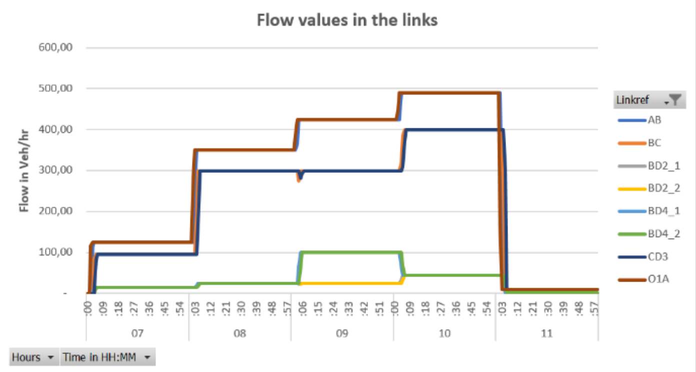
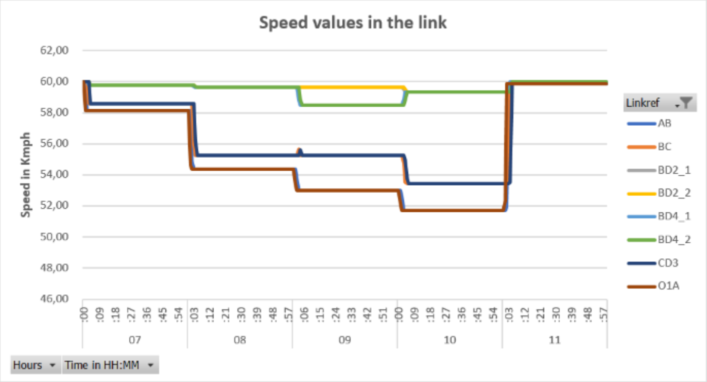


Figure 131: Turn data of the urban intersection B, throughout the network loading phase_3.2.1_MaDAM

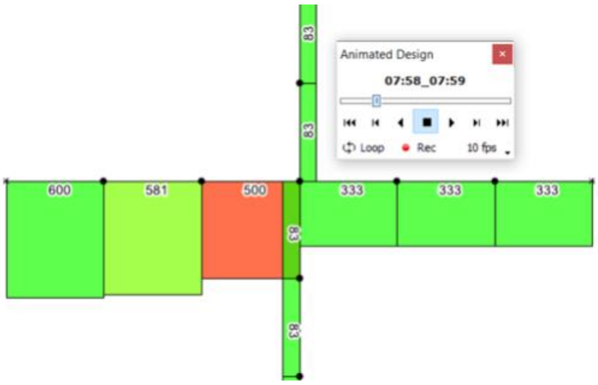
Results match expectation and propagation go uninterrupted without congestion, as it is undersaturated. As expected, the route travel time increases step by step and reaches the peak when the saturation is 98%. This suggests that only deterministic (uniform) primary delay is modeled, and no extra delay due to stochastic arrivals, in which case a sharp increase for saturations above 80% would be observed (FakhraeiRoudsari, Huang, & Tampère, 2015). The main reason why congestion doesn't occur is

</td></tr>
</table>

	because, an additional 300 veh/hr is added as the saturation flow in the intersection based on the experience with the previous test cases.
Score in EMMa	**3**
Results - MARPLE	 *Figure 132: Flow values and speed values in links of the urban intersection network_3.2.1_MARPLE* The flow values described in the Figure 132, represent the link present in the corresponding arm of the intersection B shown in the center. Saturation flow values of link -2 (controlled link) is following a weird trend, based on the demand generated from O1. Ideally the capacity of the controlled link should be 500 veh/hr, which is the case only when the demand is dropped to zero. This could be an effect of the green time signal control strategies employed for the current model run. Upon discussion with the model developer, this issue was identified and resolved for future versions of MARPLE.
Score in EMMa	**1**
Results- StreamLine: eGLTM	 *Figure 133: Flow propagation values in links of the urban intersection network_3.2.1_eGLTM*

<table>
<tr><td></td><td>

<i>Figure 134: Speed values on the urban network_3.2.1_ eGLTM</i>

Results match expectation and propagation go uninterrupted without congestion, as it is undersaturated, similar to that of MaDAM. Turn data cannot be read or visualized due to the issue stated previously.</td></tr>
<tr><td>Score in EMMa</td><td>2</td></tr>
<tr><td>Test ID</td><td>3.2.2</td></tr>
<tr><td>Test Network</td><td>Same as 3.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to understand the influence of the intersection in the route travel time when there is through traffic and diverging traffic at the intersection (oversaturated condition).</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Signalized Intersection</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 3.1.1, the turn saturation of the intersection is increased to 1300 veh/hr(based on results of previous three tests).</td></tr>
<tr><td>Demand Properties</td><td><i>Table 35: Demand Profile for Test No. 3.2.2</i>
<table>
<tr><th>Time</th><th>Demand (O1-D3)</th><th>Demand (O1-D2)</th><th>Demand (O1-D4)</th><th>Sum</th><th>Saturation %</th><th>Expectation</th></tr>
<tr><td>7-8</td><td>400</td><td>100</td><td>100</td><td>600</td><td>120%</td><td>Heavy delay</td></tr>
<tr><td>8-9</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
<tr><td>9-10</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
<tr><td>10-11</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
<tr><td>11-12</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>Links in AB stretch is expected to have queuing and congestion as a result of oversaturation. This would further increase the route cost between O1D3.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

<i>Figure 135: Flow propagation values in links of the urban intersection network_3.2.2_MaDAM</i></td></tr>
</table>

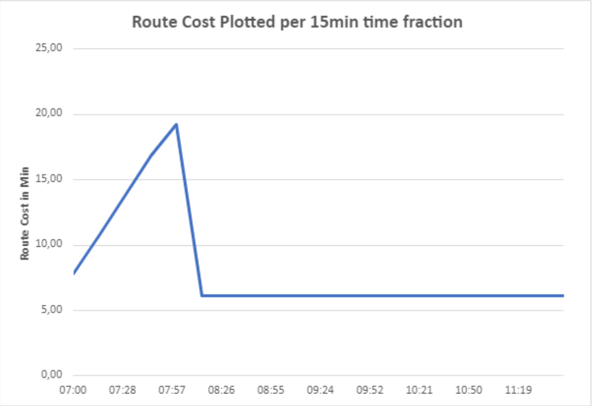


Figure 136: Travel cost in the urban road stretch AB throughout the simulation_3.2.2_MaDAM

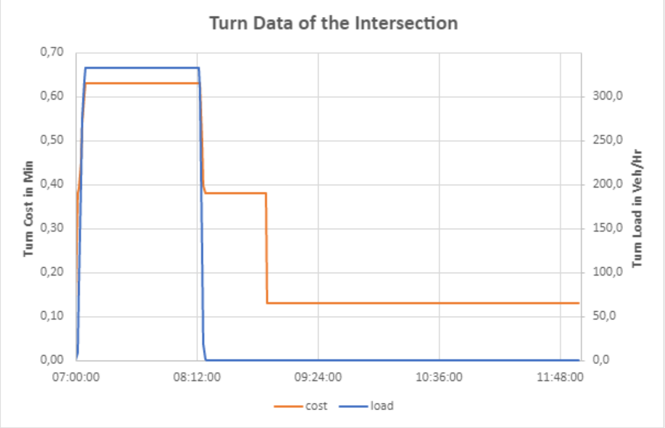


Figure 137: Turn data of the urban intersection B, throughout the network loading phase_3.2.2_MaDAM

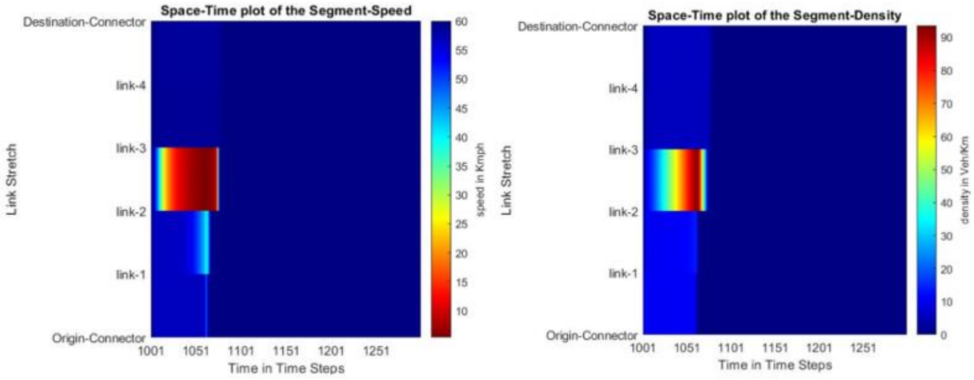
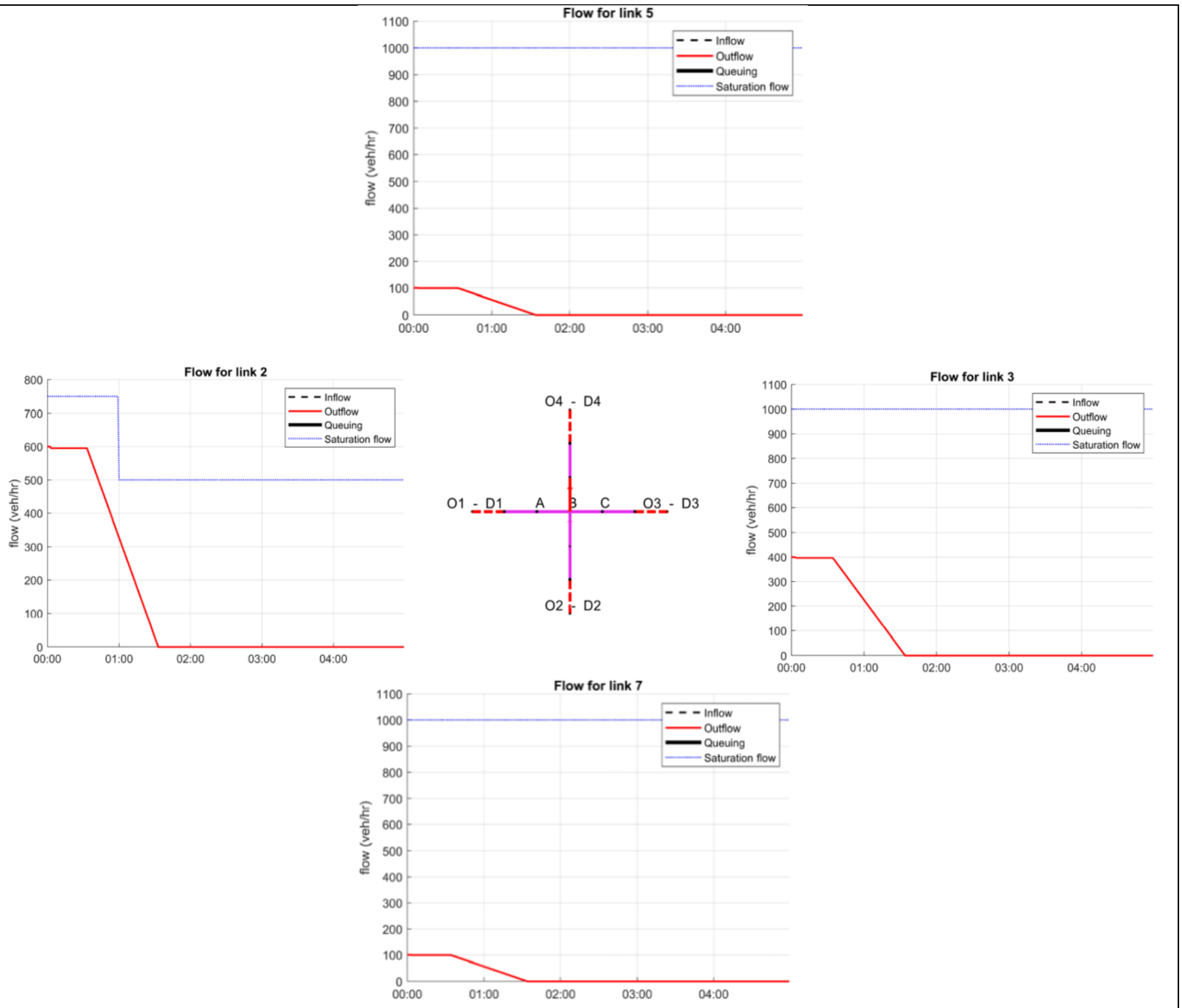
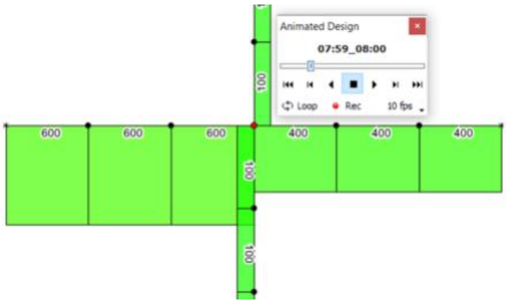


Figure 138: Space Time Diagram-Values based on Speed (Left) and Values based on Density (Right) _3.2.2_StreamLine: MaDAM

Results are as per expectation. The peak in turn load eventually leads to high turn cost as shown in Figure 136.The oversaturation in the intersection, causes a congestion in link 2, which eventually leads to a larger travel cost in the route between O1D3 as shown in Figure 137.The peak in the route cost is as a result of this congestion, till 08:15, after which the queue resolves and the network is emptied. The congestion patterns are shown in the X-T diagrams over speed and density in Figure 138. The density value of 90 Veh/km is reached, which is almost half of the jam density.

Score in EMMa	**3**

Results - MARPLE	 *Figure 139: Flow values in links of the urban intersection network_3.2.2_MARPLE* The results are not as per expectation. Due to the presence of a control link 2, the expected saturation flow should in link 2 should remain constant at 500 veh/hr, during the simulation period, a queue will be expected to form as a result of congestion (demand>capacity). But the simulation results do not provide this as shown in Figure 139. Rather the fluctuations observed in the previous test is again observed. Hence it can be concluded that the model results are not consistent when diverging links are involved.
Score in EMMa	**1**
Results- StreamLine: eGLTM	 *Figure 140: Flow propagation values in links of the urban intersection network_3.2.2_eGLTM*

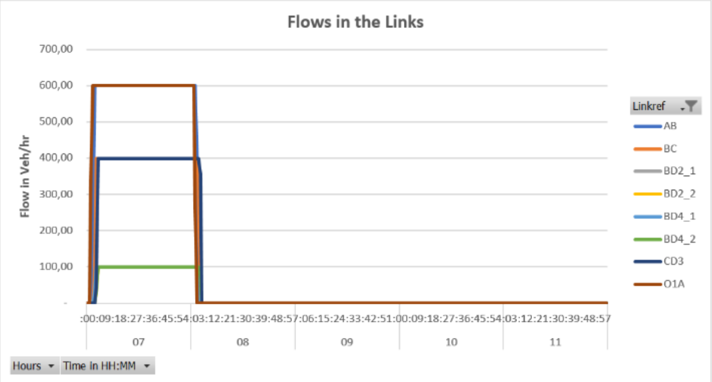
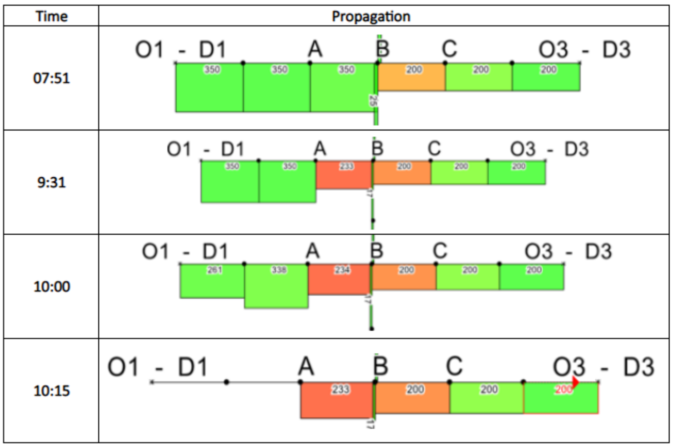
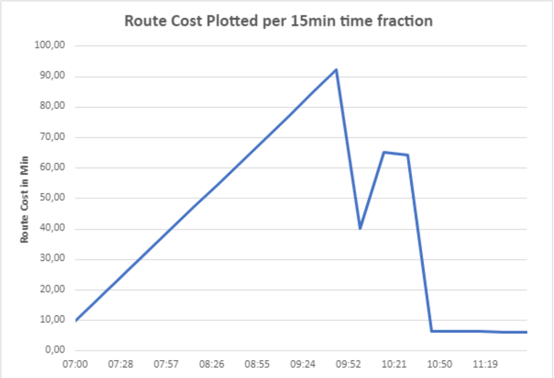
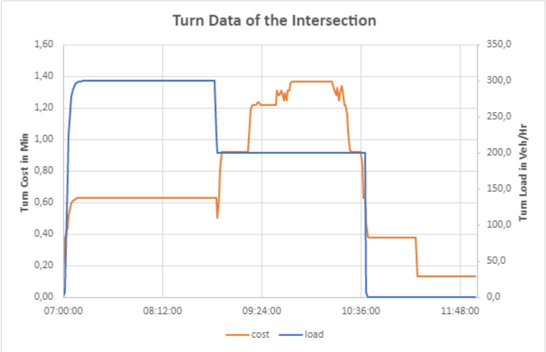
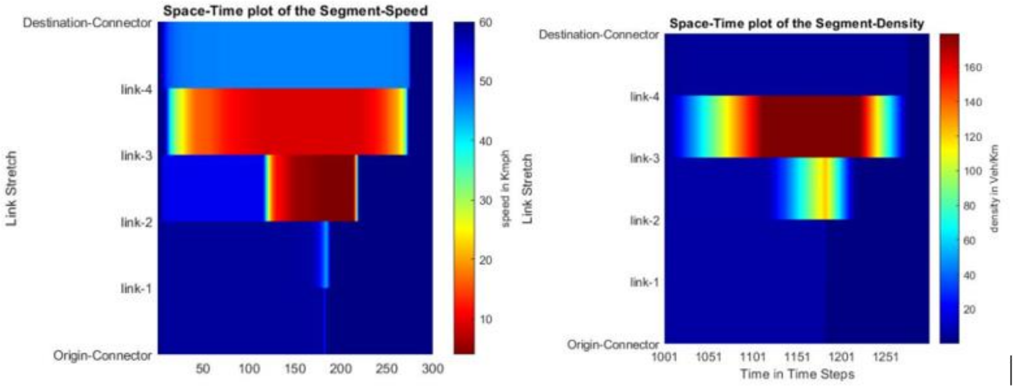
<table>
<tr><td></td><td>

Figure 141: Flow propagation values in links of the urban intersection network_3.2.2_eGLTM

Results are not in line with the expectation. The expectation here would be that the reduction in turn capacity to 50% will cause link AB to be oversaturated as the incoming demand is 600 veh/hr. However, we do not witness this reduction of capacity in AB and flow values rise beyond 500 veh/hr as shown in Figure 140 and Figure 141. This means that the reduced outflow of Link AB is more than 600 veh/hr.

To test the impact of the saturation flow of the intersection, the turn saturation value was set to 1000 veh/hr instead of 1300 veh/hr, which provided an output with congestion in link AB (results not shown here). However, in this case, the reduced outflow of the link AB due to the intersection was 525 veh/hr (quite arbitrary!). Thus, it can be concluded that the junction model in eGLTM provides inconsistent behavior.</td></tr>
<tr><td>Score in EMMa</td><td>**1**</td></tr>
<tr><td>**Test ID**</td><td>**3.2.3**</td></tr>
<tr><td>Test Network</td><td>Same as 3.1.1</td></tr>
<tr><td>Test Description</td><td>The objective of this test is to understand the influence of the intersection in the route travel time when there is through traffic and diverging traffic at the intersection (oversaturated condition due to spillback from bottleneck downstream of the intersection).</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td>*Signalized Intersection*</td></tr>
<tr><td>Supply Properties</td><td>Same as 3.1.1, the turn saturation of the intersection is increased to 1300 veh/hr (based on results of previous three tests), capacity of link in stretch CO3 is reduced to 200 veh/hr</td></tr>
<tr><td>Demand Properties</td><td>*Table 36: Demand Profile for Test No. 3.2.3*

<table>
<tr><th>Time</th><th>Demand (O1-D3)</th><th>Demand (O1-D2)</th><th>Demand (O1-D4)</th><th>Sum</th><th>Saturation %</th><th>Expectation</th></tr>
<tr><td>7-8</td><td>300</td><td>25</td><td>25</td><td>350</td><td>70%</td><td>Delay expected due to spillback</td></tr>
<tr><td>8-9</td><td>300</td><td>25</td><td>25</td><td>350</td><td>70%</td><td>Delay expected due to spillback</td></tr>
<tr><td>9-10</td><td>300</td><td>25</td><td>25</td><td>350</td><td>70%</td><td>Delay expected due to spillback</td></tr>
<tr><td>10-11</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
<tr><td>11-12</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0%</td><td></td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>The bottleneck link in CD3, is expected to create queuing in BC, which would further Links in AB stretch is expected to have queuing and congestion as a result of oversaturation. This would further increase the route cost between O1D3.</td></tr>
</table>

Results - StreamLine: MaDAM	 *Figure 142: Flow propagation values in links of the urban intersection network_3.2.3_MaDAM*  *Figure 143: Travel cost in the urban road stretch AB throughout the simulation_3.2.3_MaDAM*  *Figure 144: Turn data of the urban intersection B, throughout the network loading phase_3.2.3_MaDAM*  *Figure 145: Space Time Diagram-Values based on Speed (Left) and Values based on Density (Right) _3.2.3_StreamLine: MaDAM*

<table>
<tr>
<td></td>
<td>Results similar to test case 3.1.3. The congestion from downstream affects the turn flows and thereby the turn cost (Figure 144) and the route cost O1D3 (Figure 143), as a consequence of spillback. The anomaly observed in 3.1.3 is repeated here where unusually large value of density is observed during peak (Figure 145). Furthermore, observing the route cost vs propagation chart, it is abnormal that the route cost at 10:00 AM is lower than that at 10:15 or 10:30, even when the total no of vehicles in the network is lower. Although since all turns shared the lane, we cannot see how spillback affects traffic towards other turns. To do so, lane groups should be defined with their own signal setting. This test is not performed for the current research.</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>**3**</td>
</tr>
<tr>
<td>Results - MARPLE</td>
<td>

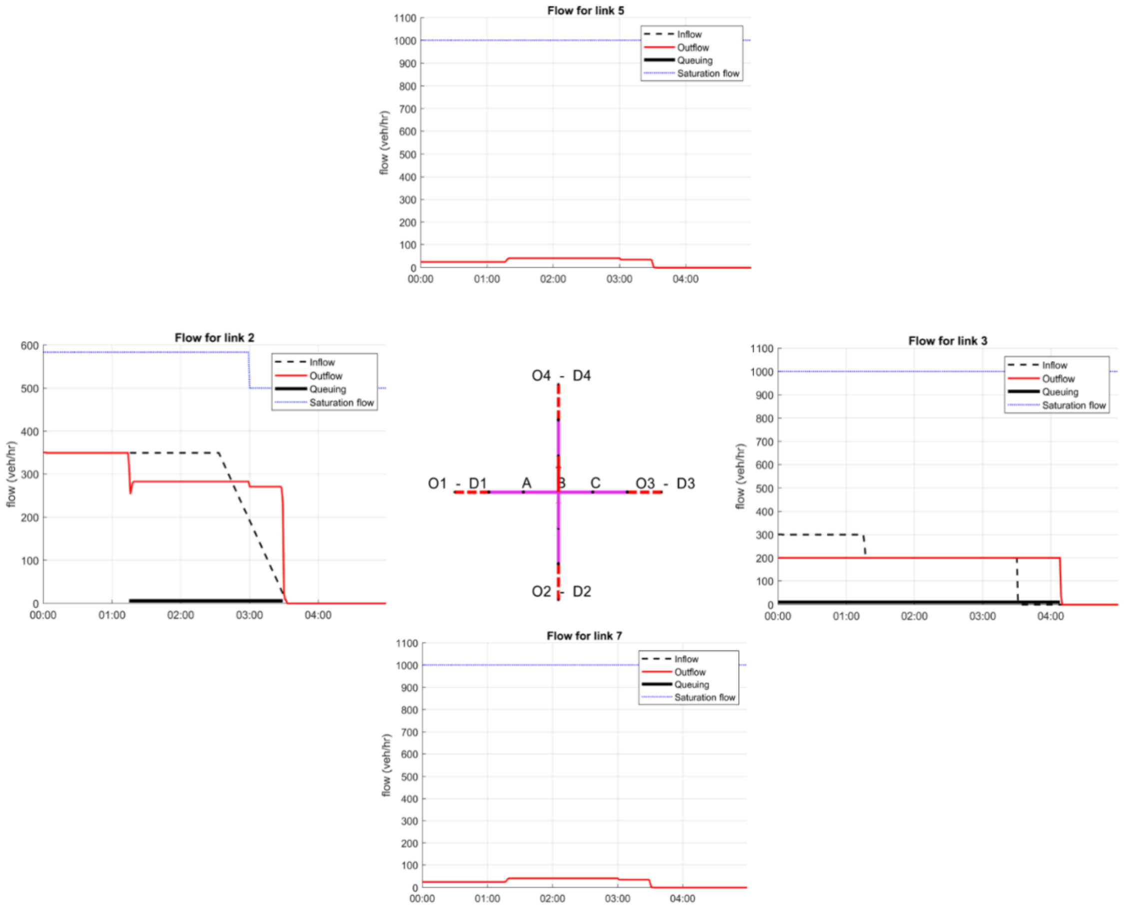


Figure 146: Flow values in links of the urban intersection network_3.2.3_MARPLE

The model results are away from the expectation. The expectation in this case would be a spillback from the downstream to the controlled link connecting the intersection (Link no-2). The capacity of the link is reduced to an arbitrary 280 veh/hr (as opposed to 200 veh/hr), which cannot be explained as an effect of the intersection control, as the saturation flow values vary over the simulation period for this link (as shown in Figure 146) . The reduction in the link capacity causes the inflow of 350 veh/hr to be blocked back and thereby create a queue and congestion in link-2. The capacity reduction can be attributed to spillback. However, the reduction here would be expected to the tune of 200 veh/hr, which is the capacity of the bottleneck link downstream of Link-2. Moreover, instances of queuing is also observed in the bottleneck link, which is brings the question of accuracy in queue formation.
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>**1**</td>
</tr>
</table>

Results-StreamLine: eGLTM

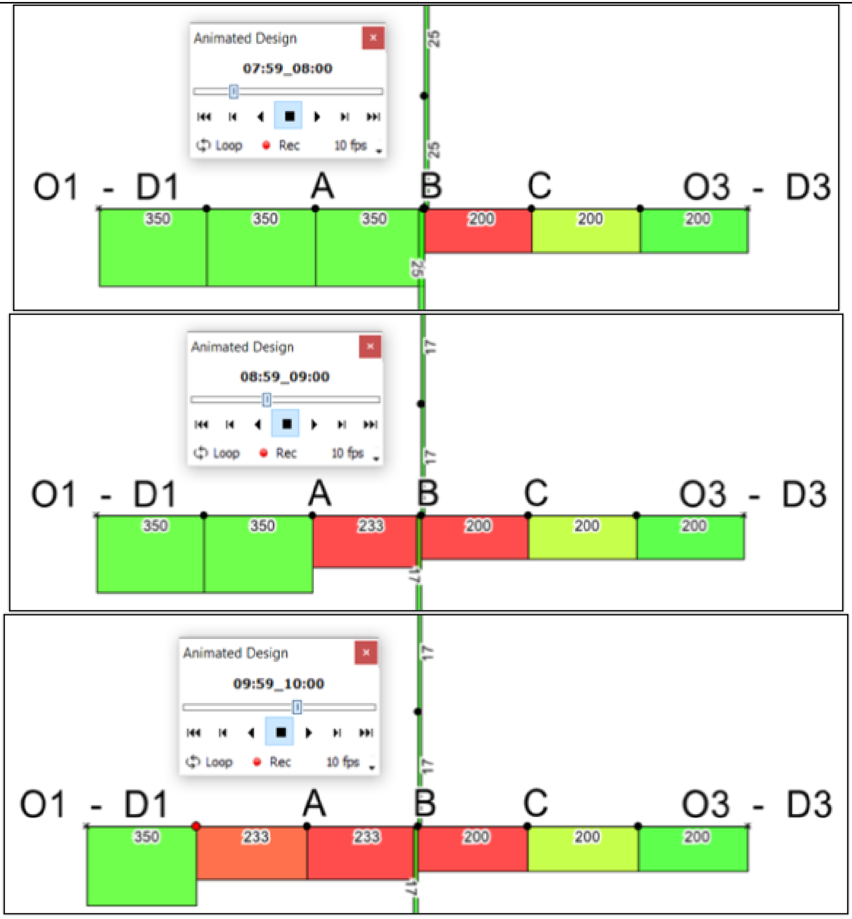


Figure 147: Flow propagation values in links of the urban intersection network_3.2.3_eGLTM

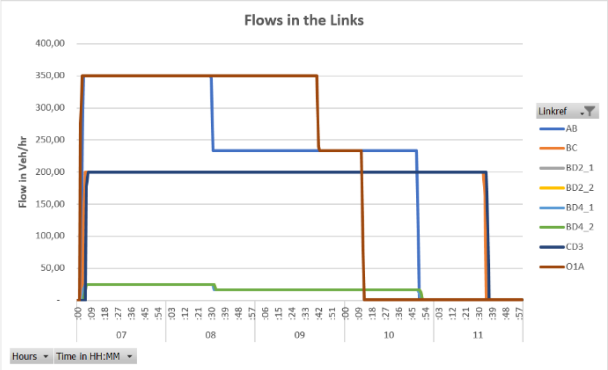


Figure 148: Flow values in links of the urban intersection network_3.2.3_eGLTM

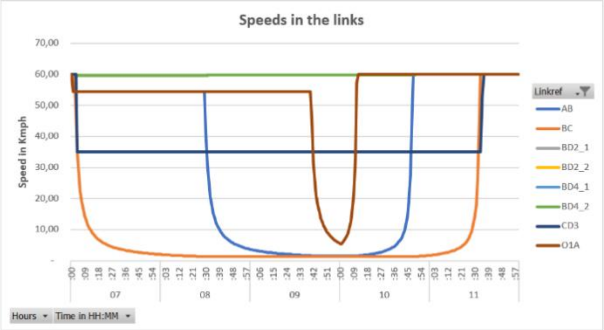


Figure 149: Speed values in links of the urban intersection network_3.2.3_eGLTM

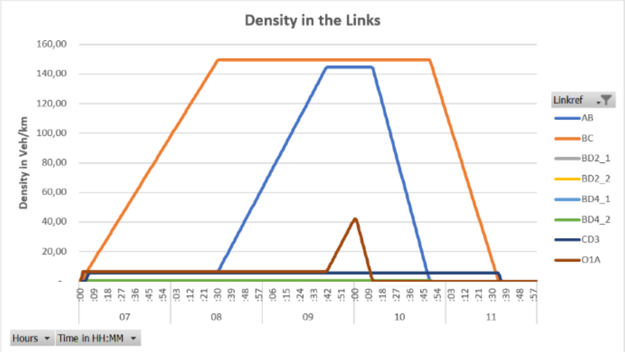
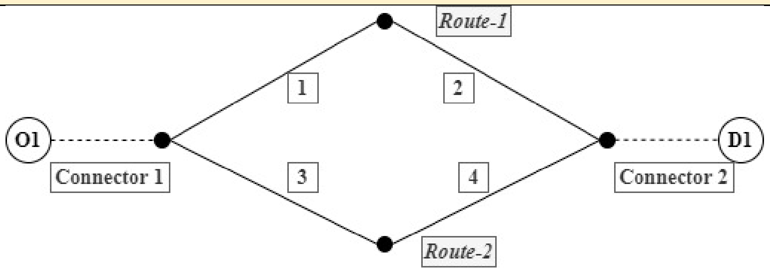
<table>
<tr><td></td><td>


Figure 150: Density values in links of the urban intersection network_3.2.3_eGLTM

Results are in line with expectation. Heavy congestion in link AB due to spillback from downstream link CD3 → BC (as shown in Figure 147). Consequent to congestion, a heavy speed drop is observed in link adjacent to bottleneck – BC (Figure 150). Although the values are abnormally low and the density values are abnormally high (150 veh/km almost close to Jam density). Node divergent constraints are strictly followed as the values in BD4 and BD2 drops, when the flow values in AB drops.
</td></tr>
<tr><td>Score in EMMa</td><td>3</td></tr>
</table>

4. Route choice submodule

The series of tests are conducted to understand the route choice behavior and the influence of delay and queuing in route travel time cost (due to feedback mechanism). The test networks mainly adopted from (FakhraeiRoudsari, Huang, & Tampère, 2015) and (Chen, Kasikitwiwat, & Ji, 2003).

<table>
<tr><td>Test ID</td><td>4.1.1</td></tr>
<tr><td>Test Network</td><td>


</td></tr>
<tr><td>Test Description</td><td>To test simple route choice behavior with all the links in undersaturated conditions.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Route choice (general)</i></td></tr>
<tr><td>Supply Properties</td><td>

<i>Table 37: Network properties for Test No. 4.1.1</i>

<table>
<tr><th></th><th>Capacity (veh/hr)</th><th>Length (Km)</th><th>Free Flow Speed (Kmph)</th><th>Speed at Cap (Kmph)</th><th>No of Lanes</th></tr>
<tr><td>Corridor link</td><td>5000</td><td>0.5</td><td>60</td><td>30</td><td>1</td></tr>
<tr><td>Other links</td><td>2000</td><td>0.5</td><td>60</td><td>40</td><td>1</td></tr>
</table>

<table>
<tr><td>Route Choice Averaging (SUE)</td><td>MSA</td></tr>
<tr><td>Initial Route Choice (SUE)</td><td>MNL</td></tr>
<tr><td>Pre-trip route Choice (SUE)</td><td>MNL</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td>Demand Properties</td><td>

Table 38: Demand Input for Test No. 4.1.1

<table>
<tr><td>Simulation Time</td><td>3 Hrs</td></tr>
<tr><td>Demand Profile</td><td>[1000, 1000, 0]</td></tr>
<tr><td>Default No. of Iterations</td><td>50</td></tr>
<tr><td>Duality Gap Threshold</td><td>0.01</td></tr>
<tr><td>Route Generator</td><td>MOTECARLO</td></tr>
</table>

</td></tr>
<tr><td>Expectation</td><td>As route costs of both the routes are similar, we expect equal route proportion throughout the simulation period.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

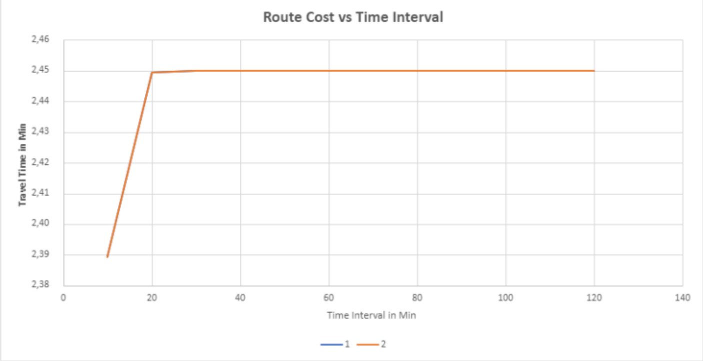


Figure 151: Travel time cost of the routes_4.1.1_MaDAM

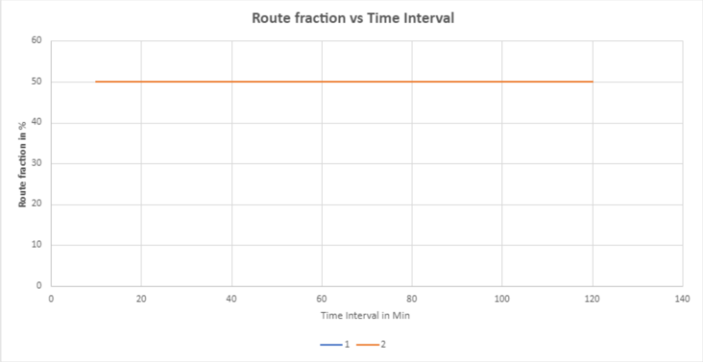


Figure 152: Route proportions per 10min interval time_4.1.1_MaDAM

Results are straightforward and in par with the expectation. The route fractions are equally split between the routes, in line with the route costs, which matches exactly with each other as shown in Figure 151 and Figure 152.

</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Results - MARPLE</td><td>

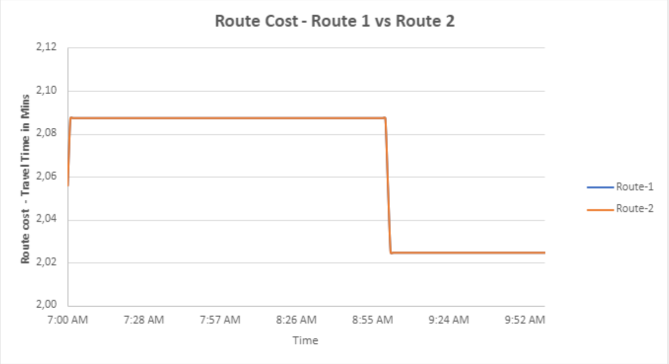
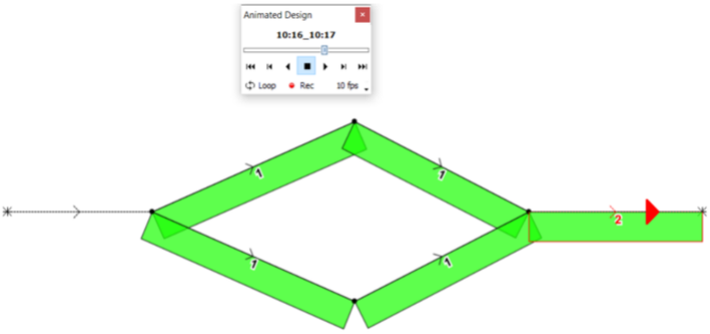
Table 39: Route flows per time period_4.1.1_MARPLE

<table>
<tr><th>Route flows in Veh/hr</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th></tr>
<tr><td>Route-1</td><td>500</td><td>500</td><td>500</td></tr>
<tr><td>Route-2</td><td>500</td><td>500</td><td>500</td></tr>
</table>

Table 40:Route travel cost per time period_4.1.1_MARPLE

<table>
<tr><th>Route Cost in Min</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th></tr>
<tr><td>Route-1</td><td>2</td><td>2</td><td>2</td></tr>
<tr><td>Route-2</td><td>2</td><td>2</td><td>2</td></tr>
</table>

</td></tr>
</table>

	Results are in line with the expectation, as shown in Table 39 and Table 40. Route cost calculated are based on free speed as per expectation, as the links are in undersaturated conditions.
Score in EMMa	**4**
Results-StreamLine: eGLTM	 *Figure 153: Travel time cost of the routes_4.1.1_eGLTM*  *Figure 154: Flow propagation chart of the network_4.1.1_eGLTM* Results are straightforward and in par with the expectation. The route fractions are equally split between the routes, in line with the route costs, which matches exactly with each other as shown in Figure 153. Abnormality is when emptying the network where a value of 1veh/hr is still present in all the links of the network, which is absurd as shown in Figure 154.
Score in EMMa	**3**
Test ID	**4.1.2**
Test Network	Same as 4.1.1
Test Description	To test simple route choice behavior with all the links in undersaturated conditions, when cost of one route slightly more than other
MoP Evaluated - Link to EMMa	*Route choice (general)*
Supply Properties	Same as 4.1.1, Except the length of the routes, as provided below: Total length of route-1 = 21 Kms Total length of route-2 = 21.5 Kms

<table>
<tr>
<td>Demand Properties</td>
<td>

Table 41: Demand Input for Test No. 4.2.1

<table>
<tr><td>Simulation Time</td><td>3 Hrs</td></tr>
<tr><td>Demand Profile</td><td>[1000, 2000, 0]</td></tr>
<tr><td>Default No. of Iterations</td><td>50</td></tr>
<tr><td>Duality Gap Threshold</td><td>0.01</td></tr>
<tr><td>Route Generator</td><td>MOTECARLO</td></tr>
</table>

</td>
</tr>
<tr>
<td>Expectation</td>
<td>Larger route fraction and flows will be expected for the less costlier route (route-1) as a result of lesser travel time.</td>
</tr>
<tr>
<td>Results - StreamLine: MaDAM</td>
<td>

The initial run of the test involved the speed values of the links in route-2 to drop abnormally (results not shown here). A possible explanation of this could be that the high density in the connecting link to the origin, is misinterpreted as a queue in the southern route (route-2) and the traffic is made to anticipated for lower speed. The speed drop also resulted in substantial increase of route cost. This shows a fallout of the second-order model.

To remove this anomaly, the receiving link's length was increased and the tests were performed again, which resulted in removal of this speed drop. The results are thus plotted for this corrected network as shown in Figure 155 and Figure 156. As expected, the shorter route (route-1), had the majority of the load in comparison.

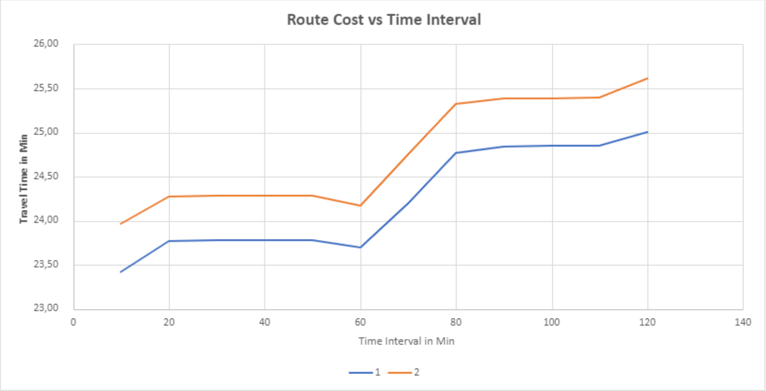


Figure 155: Travel time cost of the routes_4.1.2_MaDAM

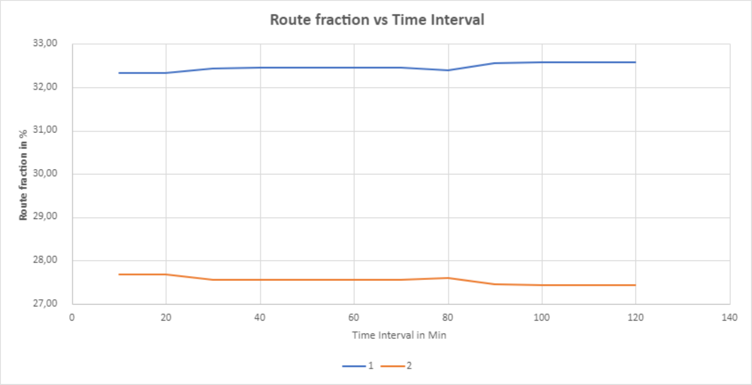


Figure 156: Route proportions per 10min interval time_4.1.2_MaDAM

</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>3</td>
</tr>
</table>

<table>
<tr><td>Results - MARPLE</td><td>

Table 42: Route flows per time period_4.1.2_MARPLE

<table>
<tr><th>Route flows in Veh/hr</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th></tr>
<tr><td>Route-1</td><td>1241.89</td><td>621.022</td><td>0</td></tr>
<tr><td>Route-2</td><td>758.106</td><td>378.978</td><td>0</td></tr>
</table>

Table 43: Route travel cost per time period_4.1.2_MARPLE

<table>
<tr><th>Route Cost in Min</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th></tr>
<tr><td>Route-1</td><td>1.54</td><td>1.54</td><td>1.53</td></tr>
<tr><td>Route-2</td><td>2.03</td><td>2.03</td><td>2.03</td></tr>
</table>

As expected, the shorter route, had the majority of the load in comparison, evident from Table 42 and Table 43.

</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
<tr><td>Results-StreamLine: eGLTM</td><td>

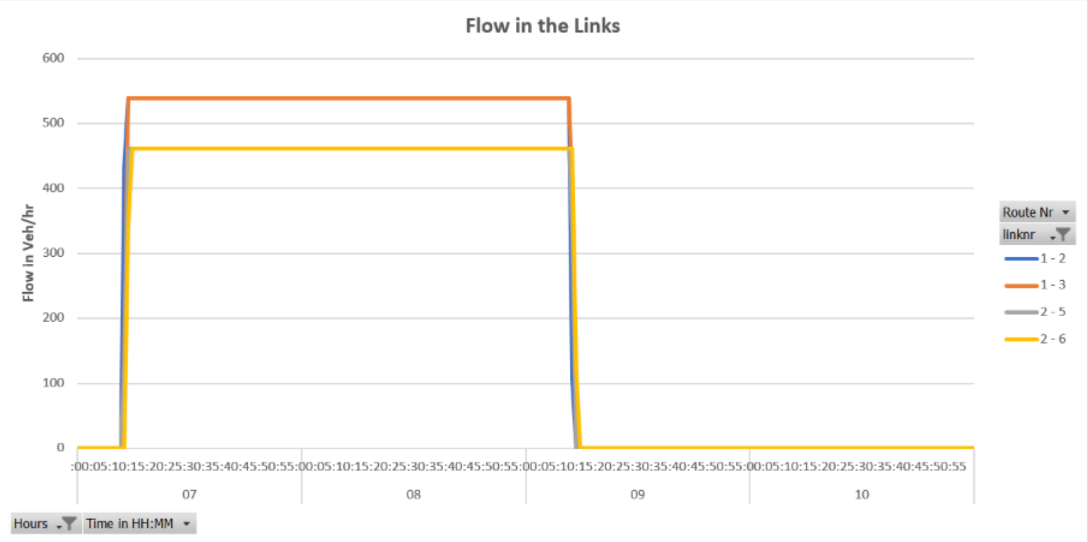


Figure 157: Route flows in the Links (legend: routenr-linknr) _4.1.2_eGLTM

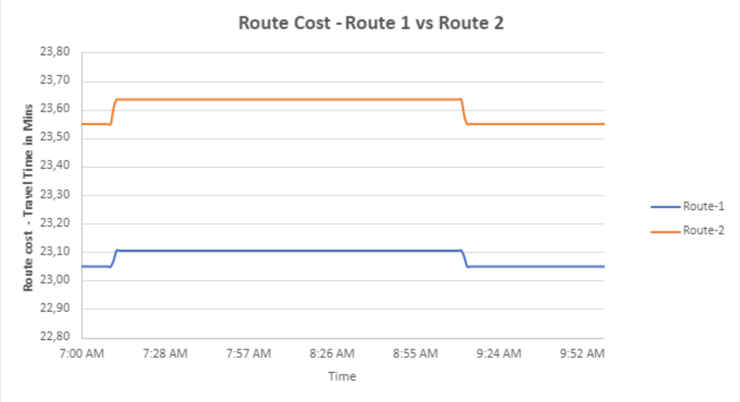


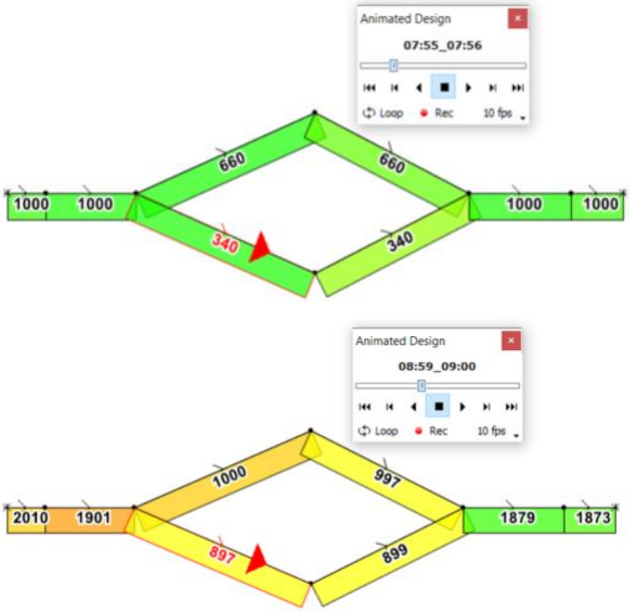
Figure 158: Route travel cost for final iteration_4.1.2_MARPLE

Clearly route 1 is cheaper than route 2 and therefore, flow is higher in route 1 as expected as shown in Figure 157 and Figure 158.

</td></tr>
<tr><td>Score in EMMa</td><td>4</td></tr>
</table>

<table>
<tr><th>Test ID</th><th>4.1.3</th></tr>
<tr><td>Test Network</td><td>Same as 4.1.1</td></tr>
<tr><td>Test Description</td><td>To test simple route choice behavior with some of the links in oversaturated conditions, when cost of one route slightly more than other. The route which is cheaper is oversatured.</td></tr>
<tr><td>MoP Evaluated - Link to EMMa</td><td><i>Route choice (general)</i></td></tr>
<tr><td>Supply Properties</td><td>Same as 4.1.1, Capacity of all links are reduced to 1000 veh/hr. The length of the routes, as provided below:

Total length of route-1 = 21 Kms

Total length of route-2 = 22 Kms</td></tr>
<tr><td>Demand Properties</td><td><i>Table 44: Demand Input for Test No. 4.3.1</i>
<table>
<tr><td>Simulation Time</td><td>3 Hrs</td></tr>
<tr><td>Demand Profile</td><td>[1000, 3000, 0]</td></tr>
<tr><td>Default No. of Iterations</td><td>50</td></tr>
<tr><td>Duality Gap Threshold</td><td>0.01</td></tr>
<tr><td>Route Generator</td><td>MOTECARLO</td></tr>
</table></td></tr>
<tr><td>Expectation</td><td>The upper route (route-1) is expected to have larger route fraction in the first hour of the simulation. However, during the second hour of the simulation, the travel time of route-1 is expected to increase and the route fraction would incline more towards the southern, now cheaper route, till the time vehicles achieve free-flow speed, which is when the oversaturation is removed.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>


<i>Figure 159: Flow propagation chart of the network_4.1.3_MaDAM</i></td></tr>
</table>

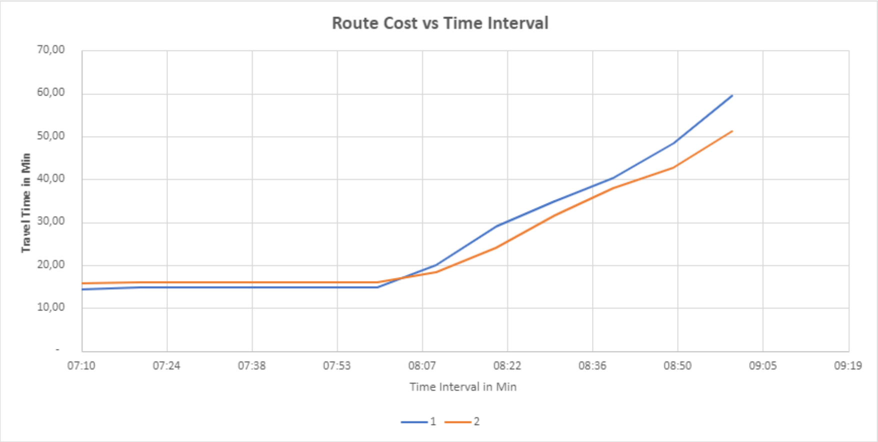


Figure 160: Travel time cost of the routes per 10min interval_4.1.3_MaDAM

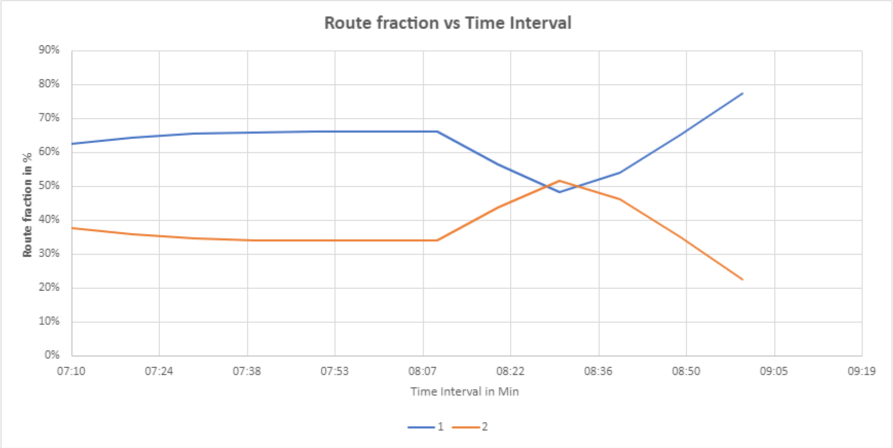


Figure 161: Route proportions per 10min interval time_4.1.3_MaDAM

Results are in line with expectation. The upper route (route-1) attracts more flows in the first hour of the simulation, after which as a result of oversaturation and congestion in the upper route, queuing occurs and thereby travel cost increases as shown in Figure 160. A spillback can also be observed in the link upstream of the congestion as shown in Figure 159. The route fraction drops for the upper route as a result of the travel time, however it recovers back when all the links become equally saturated and the upper route again becomes more attractive (ref to Figure 161).

Score in EMMa

4

Results - MARPLE

Table 45: Route flows per time period_4.1.3_MARPLE

Route flows in Veh/hr	Time period-1	Time period-2	Time period-3
Route-1	727.366	1144.621	0
Route-2	272.634	1855.379	0

Table 46: Route travel cost per time period_4.1.3_MARPLE

Route Cost in Min	Time period-1	Time period-2	Time period-3
Route-1	14.07	24.04	21.81
Route-2	15.05	23.55	20.97

<table>
<tr>
<td></td>
<td>

Links of Northern Route No. 1

Links of Southern Route No. 2

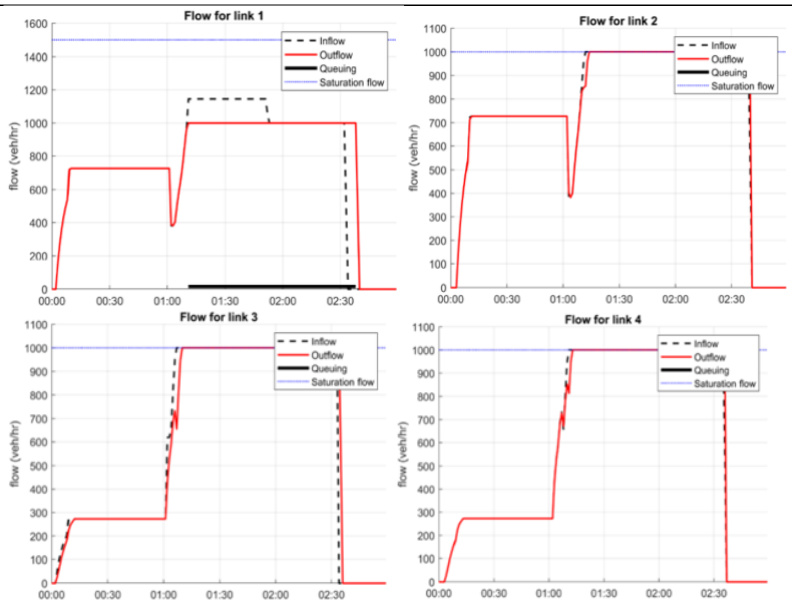


Figure 162: Flow values in the links of the network_4.1.3_MARPLE

Table 47: Route delays per time period_4.1.3_MARPLE

<table>
<tr><th>Route Delay Time in Min</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th></tr>
<tr><td>Route-1</td><td>0.07</td><td>10.04</td><td>7.81</td></tr>
<tr><td>Route-2</td><td>0.05</td><td>8.55</td><td>5.97</td></tr>
</table>

The results are as per expectation. During the first hour of the simulation, the flow is higher on the upper route compare to the lower route as the upper route is cheaper to travel. This is effectively captured by the model with the route fractions and the flow values during the beginning simulation time periods, as shown in Table 45 and Table 46. In the second hour of the simulation, a queue formation starts in the link-1 (Route-1) as a result of blocking back by link-2 (Route-1) as a result of flow>capacity. As a result of oversaturation, the route cost gets higher due to larger delays in the upper route (rRoute-1), which can be observed from the delay values per time period as shown in Table 47. This has resulted in a dip in flow values in the link flow graphs represented for the upper route links as shown in Figure 162. Overall, the model behavior is as per expectation.

</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>4</td>
</tr>
<tr>
<td>Results-StreamLine: eGLTM</td>
<td>

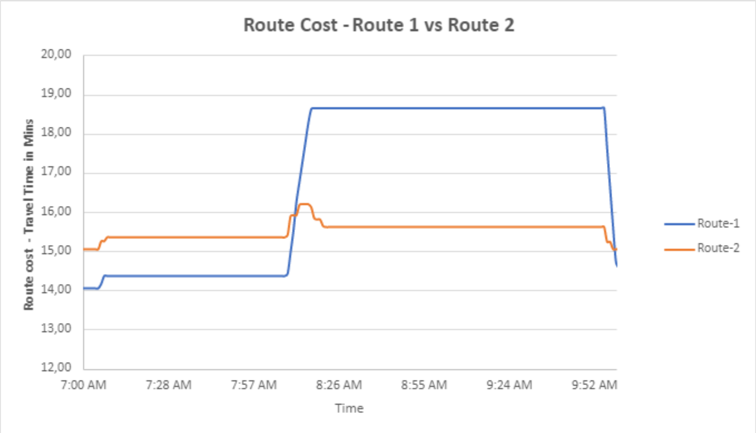
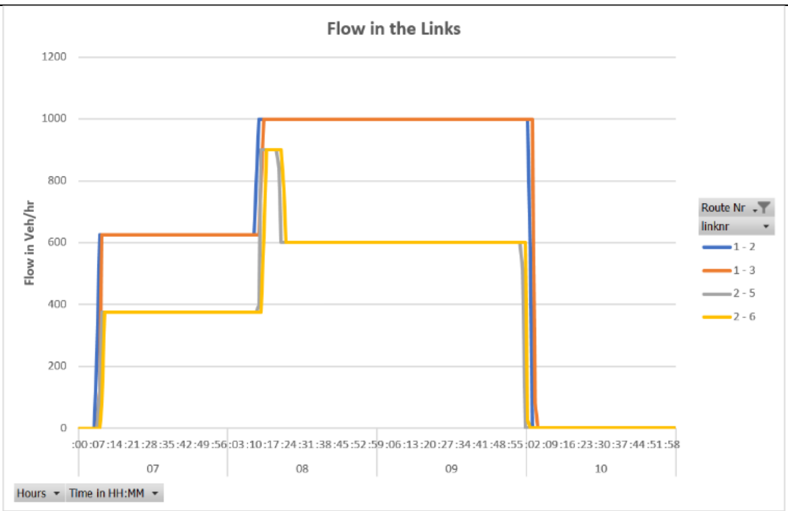
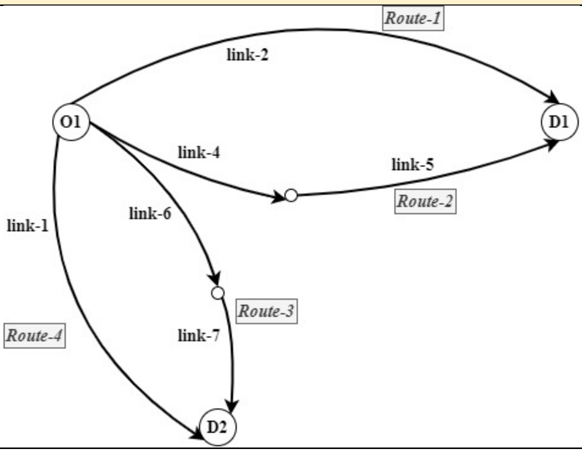


Figure 163: Travel time cost of the routes_4.1.3_eGLTM

</td>
</tr>
</table>

	 *Figure 164: Route flows in the Links (legend: routenr-linknr) _4.1.3_eGLTM* Results are in line with expectation. The upper route attracts more flows in the first hour of the simulation, after which as a result of oversaturation and congestion in the upper route, queuing occurs and thereby travel cost increases as shown in Figure 163. Route cost graphs are a clear indication of how route one becomes substantially costlier due to the delay (consequence of oversaturation). The chart describing link flows in Figure 164, shows the influence of higher route costs in route-1 to divert the traffic to route-2. This can be understood by sudden rise in flows in links 5 and 6 of route 2, at 08:00 AM.
Score in EMMa	**4**
Test ID	**4.2.1**
Test Network	
Test Description	A four-route network of independent routes is tested, to understand the general route choice behavior.
MoP Evaluated - Link to EMMa	*Route choice (general)*

<table>
<tr><td>Supply Properties</td><td>

Table 48: Network properties for Test No. 4.2.1

<table>
<tr><th>Link No</th><th>Route No</th><th>Length (Km)</th><th>Free Speed (Kmph)</th><th>Capacity (veh/h)</th></tr>
<tr><td>1</td><td>4</td><td>5</td><td>100</td><td>2000</td></tr>
<tr><td>2</td><td>1</td><td>10</td><td>100</td><td>2000</td></tr>
<tr><td>4</td><td>2</td><td>4</td><td>120</td><td>2000</td></tr>
<tr><td>5</td><td>2</td><td>6</td><td>120</td><td>1000</td></tr>
<tr><td>6</td><td>3</td><td>4</td><td>120</td><td>2000</td></tr>
<tr><td>7</td><td>3</td><td>1</td><td>120</td><td>1500</td></tr>
</table>

Route Choice Averaging (SUE) :MSA
Initial Route Choice (SUE) :MNL
Pre-trip route Choice (SUE) :MNL

Length of the Route 1 and Route 2 : 10Kms
Length of the Route 3 and Route 4 : 5Kms
</td></tr>
<tr><td>Demand Properties</td><td>

Table 49: Demand Profile for Test No. 4.2.1

<table>
<tr><th>Demand in Veh/hr</th><th>07-08</th><th>08-09</th><th>09-10</th><th>10-11</th></tr>
<tr><td>O1-D1</td><td>2000</td><td>2000</td><td>0</td><td>0</td></tr>
<tr><td>O1-D2</td><td>2000</td><td>2000</td><td>0</td><td>0</td></tr>
</table>
</td></tr>
<tr><td>Expectation</td><td>Unrestricted flow of demand is expected in all the links. As for a given OD pair, the cost of each route is same (length of the route is equal), an equal or similar route proportioning is expected.</td></tr>
<tr><td>Results - StreamLine: MaDAM</td><td>

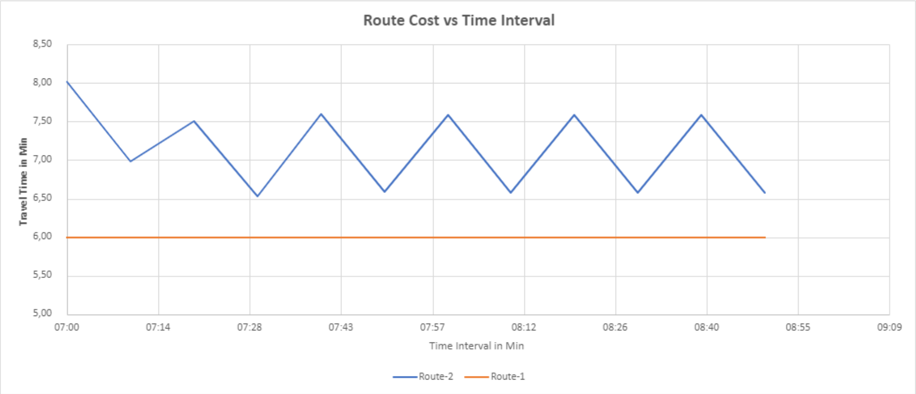


Figure 165: Travel time cost of the routes_O1D1_4.2.1_MaDAM

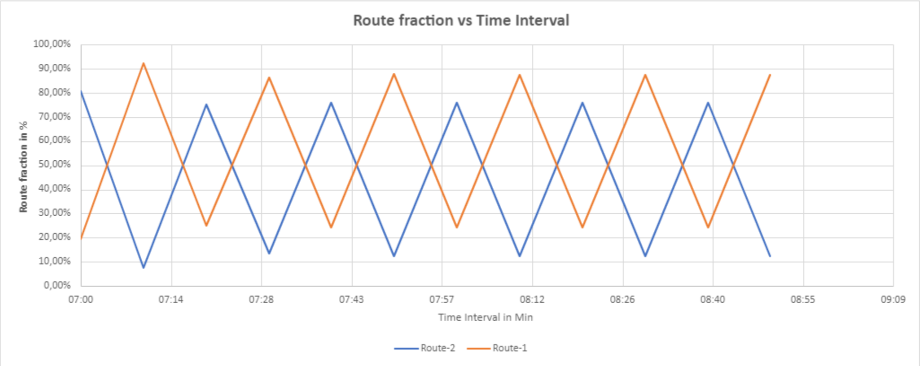
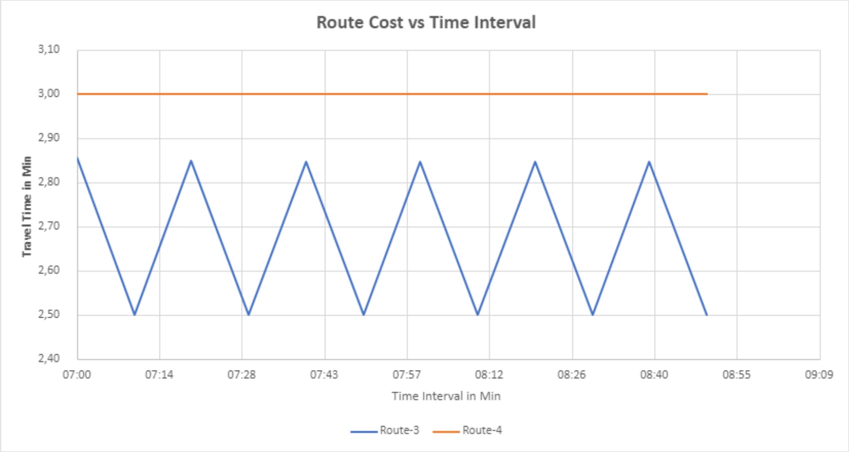
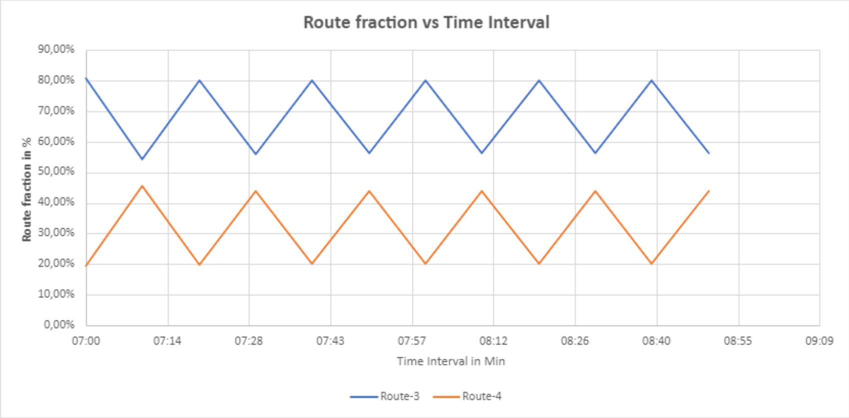


Figure 166: Route proportions per 10min interval time_O1D1_4.2.1_MaDAM
</td></tr>
</table>

	The results are not in line with the theoretical expectation for the centroids O1-D1(centroids: 1-2). This is can be observed from the results of the route fractions and the route costs in Figure 165 and Figure 166. For O1D1, we observe that the route cost calculation begins with a significantly higher cost for route-2 compared route-1, even though length of both routes is same. However, the route proportions obtained is not logical. Considering the route costs as per Figure 165, we would expect a higher route fraction for the route with lower cost. For the lower route 2, the route fraction fluctuated between 80% and 10% approximately, even though at all route choice time periods, the upper route 1, cost is lower. The fluctuations in the route cost can be understood as a direct consequence of oversaturation in Link-5, as the link capacity is restricted to 1000 veh/hr. The larger route proportion to the costlier route might be a consequence of the error term in the utility function, but it cannot be found realistic. Interestingly, For the O1-D2 pair (centroids: 1-2), the route choice proportions for the costlier route are lower and consistent with the expectation as shown in Figure 167 and Figure 168.  *Figure 167: Travel time cost of the routes_O1D2_4.2.1_MaDAM*  *Figure 168: Route proportions per 10min interval time_O1D2_4.2.1_MaDAM* Total Simulation time: 145.8919 seconds Relative duality gap value between final two iterations = 0.008782 Peak Memory Usage of the final iteration = 298 MBs
Score in EMMa	**1**

Results - MARPLE

Table 50: Route flows per time period_4.2.1_MARPLE

OD Pair	**Route flows in Veh/hr**	**Time period-1**	**Time period-2**	**Time period-3**	**Time period-4**
O1D1	Route-1	1001.331	1003	0	0
O1D1	Route-2	998.669	997.1	0	0
O1D2	Route-3	1239.868	1239	0	0
O1D2	Route-4	760.132	761.2	0	0

Table 51: Route travel cost per time period_4.2.1_MARPLE

OD Pair	**Route Cost in Min**	**Time period-1**	**Time period-2**	**Time period-3**	**Time period-4**
O1D1	Route-1	6.01	6.01	6	6
O1D1	Route-2	6.01	6.02	5	5
O1D2	Route-3	2.51	2.52	2.5	2.5
O1D2	Route-4	3	3	3	3

Results are in line with expectation. For O1D1, both the routes have similar route costs (Table 51), due to which the loading in both routes are almost equal evident from the route flow table (Table 50). For O1D2, Route 3 is cheaper than route 4 due to which it attracts larger load matching the expectation.

Total Simulation time: 3.74 seconds
Relative duality gap value between final two iterations = 0.008782
Peak Memory Usage of the final iteration = 0.0468 MBs

Score in EMMa

4

Results- StreamLine: eGLTM

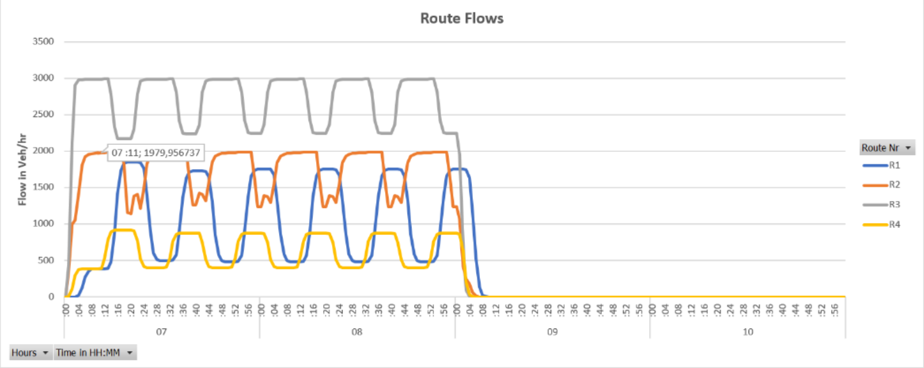


Figure 169: Route Flows_4.2.1_eGLTM

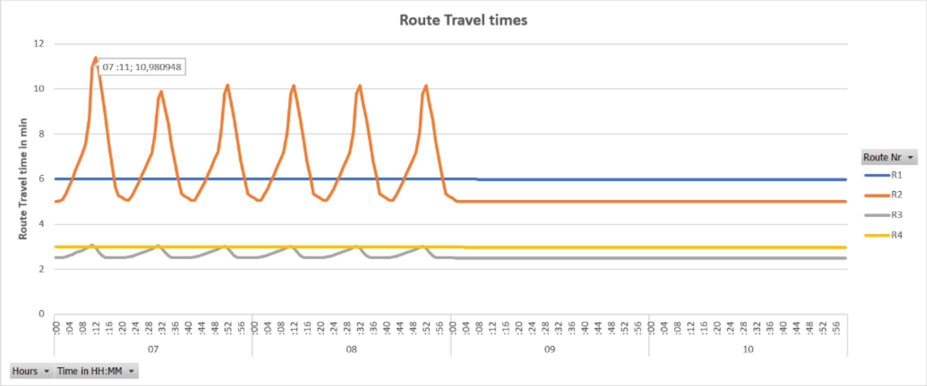
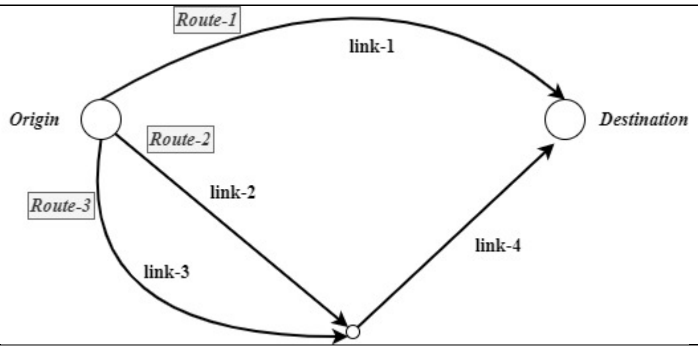


Figure 170: Travel time cost of the routes_4.2.1_eGLTM

The results are not in line with the theoretical expectation for the centroids O1-D1. This is can be observed from the results of the route flows (Figure 169) and the route costs (Figure 170). For a SUE with MNL route proportioning, we would expect a higher route fraction for the route with lower cost. At the time step 07:11, we observe that even though the route cost for R2 is higher than R1, larger flows are observed in R2, which means the route has attracted more traffic which is absurd, considering the high travel cost. The fluctuations in the route cost for R2 can be understood as a direct consequence of oversaturation in Link-5 (StreamLine Link nr-15), as the link capacity is restricted to 1000 veh/hr. Similar to MaDAM, for the O1-D2 pair the route choice proportions for the costlier route is lower and consistent with the expectation.

Score in EMMa	1
Test ID	**4.3.1**

Test Network



Test Description	The objective of this test is to check the influence of independent routes and route overlap in route choice.
MoP Evaluated - Link to EMMa	*Route choice (route overlap)*

Supply Properties

Table 52: Network properties for Test No. 4.3.1

Link No	**Route No**	**Length (Km)**	**Capacity (veh/h)**	**Free Speed (Kmph)**	**Speed at Cap (Kmph)**
1	1	10	1500	60	40
2	2	7	1500	60	40
3	3	7	1500	60	40
4	2 & 3	3	1500	60	40

Demand Properties

Table 53: Demand profile for Test No. 4.2.1

Demand in Veh/hr	**07-08**	**08-09**	**09-10**	**10-11**
O-D	2000	2000	0	0

Expectation

In this case, link-4 is overlapping for the route2 and route3. Expectation here would be a relative comparison of route proportions of the 3 routes, for MNL and PCL/C-Logit. In PCL/C-Logit, which accounts for route overlap, we expect, a higher route preference for route-1 and thereby corresponding lower route proportions in the routes-2 and 3. This expectation is based on (Bliemer & Bovy, 2008). They consider a Monte Carlo simulation of route choices from a grid network to illustrate that routes with the same length are not equally preferred, as the more the routes overlap, the lower their probability in favor of more independent alternatives.

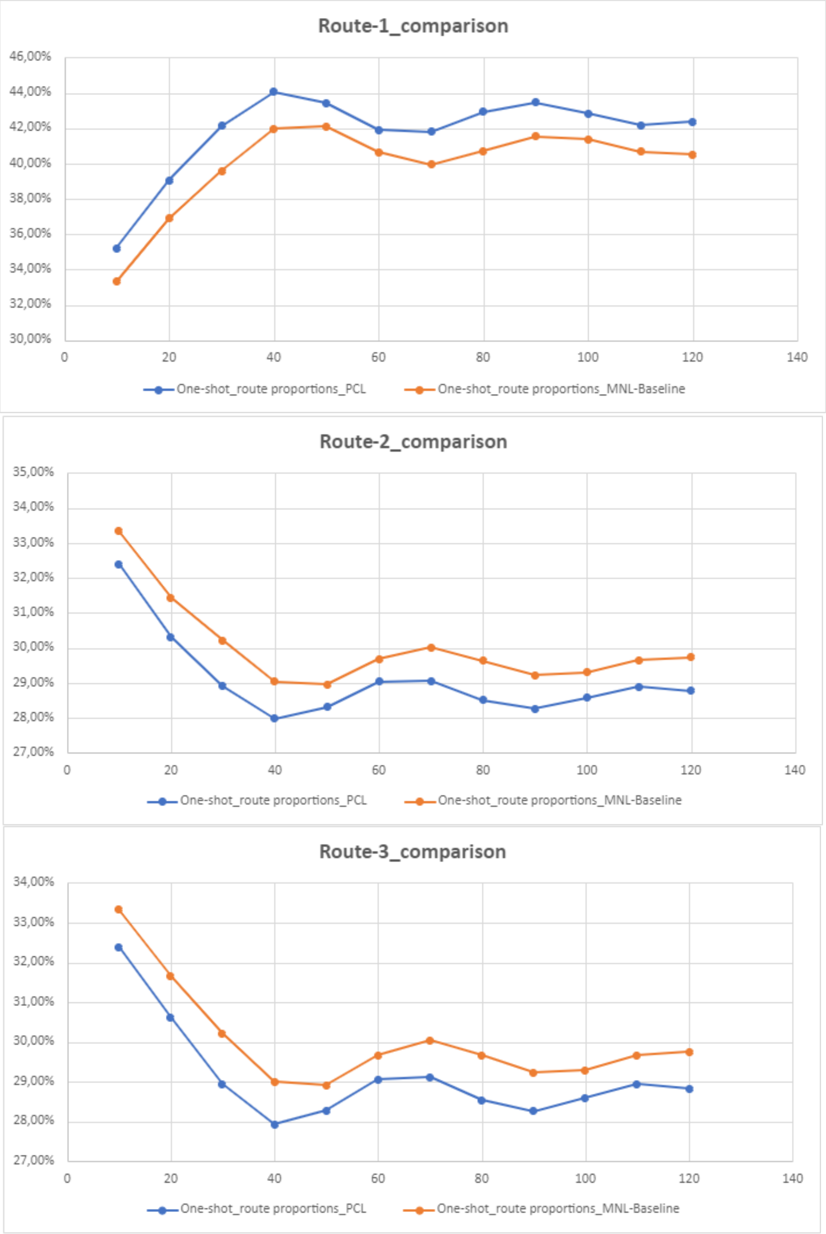
<table>
<tr>
<td>Results - StreamLine: MaDAM</td>
<td>


Figure 171: Comparison of route choice proportions Route 1(top), Route 2(middle) and Route 3(Bottom)_ 4.1.3_MaDAM

As expected, the route fraction of the upper Route 1 (independent) has increased in the PCL variant shown in Figure 171. This is an effect of route overlap in the southern routes 2 and 3 which shares the link nr 4. In the MNL model, we see almost similar route proportions for all the three routes, which is different in case of the PCL, in line with the expectations.
</td>
</tr>
<tr>
<td>Score in EMMa</td>
<td>**4**</td>
</tr>
<tr>
<td>Results - MARPLE</td>
<td>
Table 54:Route flows per time period_4.3.1_MARPLE
<table>
<tr><th>Route flows in Veh/hr</th><th>Time period-1</th><th>Time period-2</th><th>Time period-3</th><th>Time period-4</th></tr>
<tr><td>Route-1</td><td>758.298</td><td>760.025</td><td>0</td><td>0</td></tr>
<tr><td>Route-2</td><td>621.005</td><td>619.986</td><td>0</td><td>0</td></tr>
<tr><td>Route-3</td><td>620.696</td><td>619.988</td><td>0</td><td>0</td></tr>
</table>
</td>
</tr>
</table>

Table 55: Route costs per time period_4.3.1_MARPLE

Route Cost in Min	**Time period-1**	**Time period-2**	**Time period-3**	**Time period-4**
Route-1	12.09	12.10	12.01	12.00
Route-2	12.16	12.17	12.00	12.00
Route-3	12.16	12.17	12.00	12.00

MARPLE in SUE, by default uses C-Logit for the dynamic assignment which takes into account for route overlap. This is evident from the route flows (Table 54), which shows that the majority share has gone to the upper route1 compared to route 2 and route 3. This further reinforced by the fact that the route costs of all the routes are similar as shown in Table 55. The Results are in line with the expectation.

Score in EMMa

4

Results-StreamLine: eGLTM

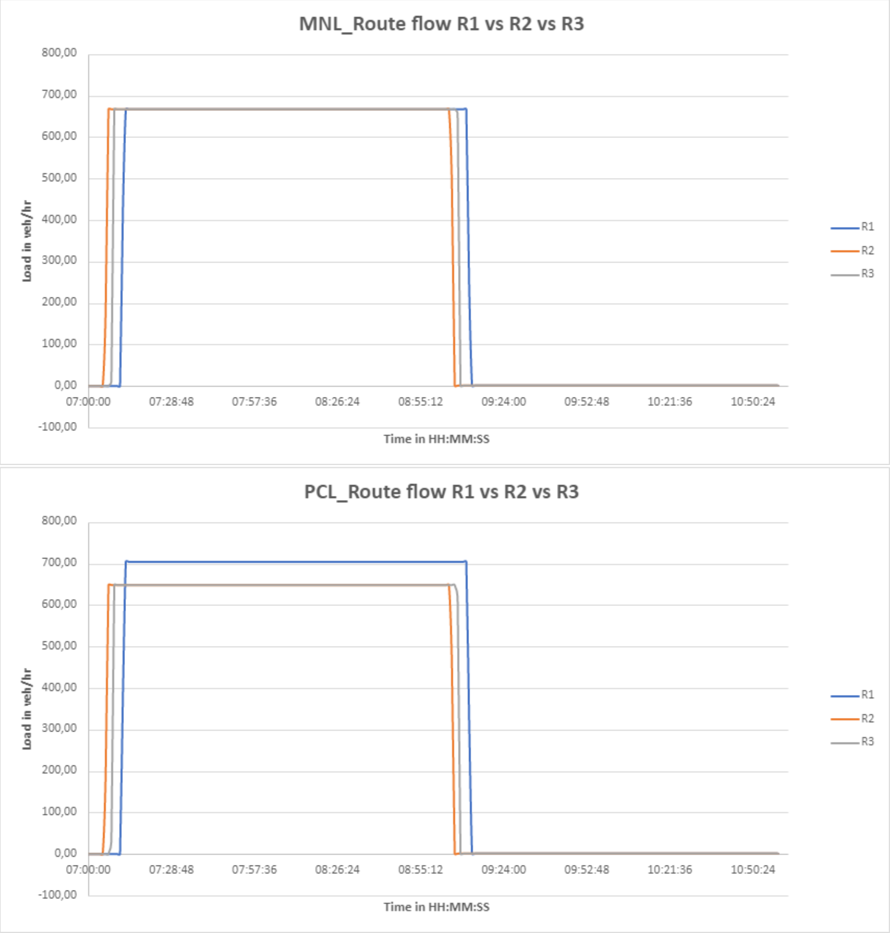


Figure 172: Route Flows MNL Variant (Upper) vs PCL Variant (Lower)_4.3.1_eGLTM

As expected, the route fraction of the upper route (independent) has increased in the PCL variant, observed from the increase in load values compared to R2 and R3 as shown in Figure 172. This is an effect of route overlap in the southern routes which shares the link nr 4. In the MNL model, we see almost similar route proportions for all the three routes, which is different in case of the PCL, in line with our expectations. Interestingly enough this becomes the case even though the route cost is increasing for R1 in the PCL variant (results not shown here) which might be a direct consequence of saturation→increasing density→decreasing speed →increasing travel time.

Score in EMMa

4

Appendix-B: Model User Survey

*Table 56: Link between the questions in the model user survey and the MoPs in EMMa for the Model Users - **Policy Maker and Mobility Consultant***

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
1	While using the transport model for the applications in the options below, how important would it be to incorporate the following real-world effect: <u>blockage of vehicles in a road section as a result of traffic jams and induced congestion in preceding road sections upstream from the blockage?</u> *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Flow Metering - strict capacity constraint,* *Traffic Spillback - strict storage constraint,* *Modeling of stop and go waves,* *Propagation - Link flows,* *Propagation – Queuing,* *Propagation - Effect of time variability in Fundamental Diagram,* *Node model-merge behavior,* *Node model-diverge behavior,* *Signalized Intersection,* *Fluctuation of traffic states over a series of urban and non-urban links*
2	While using the transport model for the applications in the options below, how important would it be to incorporate the following real-world effect: <u>decrease in the total capacity of the road section in a motorway when the traffic congestion is present for a period of time and dissolves?</u> The reason for the larger gap between the cars is due to driver's higher expected reaction time Capacity - the total number of vehicles the road section can accommodate in an hour *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Capacity drop*
3	While using the transport model for the applications listed in the options below, how important would it be to incorporate the following real-world effect: <u>gradual increase or decrease in speeds of vehicles as opposed to sudden variations?</u>	**Policy Maker and Mobility Consultant**	*Link-level dynamic distribution of vehicle speeds- Curve roughness factor,* *Link-level dynamic distribution of traffic flows- Curve roughness factor*

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
	The reason for a gradual variation is that the driver behavior is dependent on the movement of the vehicles in front and behind so as to accelerate or decelerate his/her vehicle. *(On a scale from 1 - least important to 10 - most important)*		
4	While using the transport model for the applications listed in the options below, how important would it be to incorporate the following real-world effect: The option for a traveler to have multiple route options that are made available on the basis of shorter travel times between two locations. *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Presence of variable route set,* *Dynamic Relative duality gap,* *Route choice (general),* *Route choice (route overlap)*
5	While using the transport model for the applications listed in the options below, how important would it be to incorporate different transport modes for the travelers and to observe different types of travel behavior for different types of trips such as for work, leisure etc. *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Difference in Network Supply based on Modes,* *Difference in Input parameters based on different trip purposes*
6	While using the transport model, how important it is to run the simulation in a normal computer and its ability to give fast results (quickness of the model run)? *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Run Time in Sec,* *Peak memory Usage in MB's*
7	While using the transport model for the applications listed in the options below, how important is the usability of the model, defined as the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component (Definition quoted from - IEEE Std.610.12-1990, referred from (Seffah, Donyaee, Kline, & Padda, 2006))? *(On a scale from 1 - least important to 10 - most important)*	**Policy Maker and Mobility Consultant**	*Familiarity,* *Simplicity,* *Navigability,* *Controllability,* *Readability,* *User guidance,* *Flexibility*

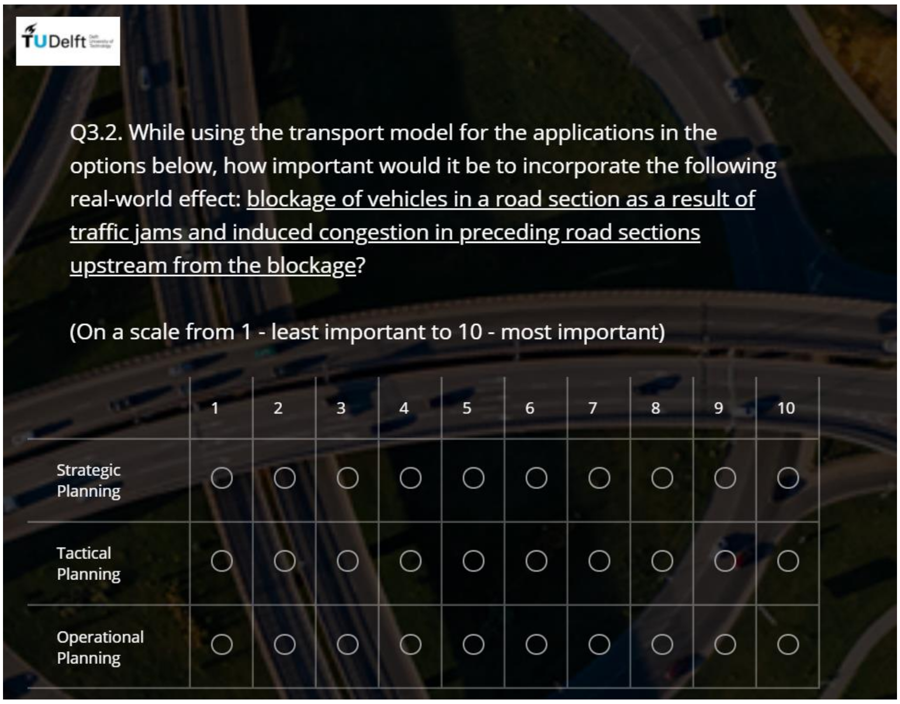
*Table 57: link between the questions in the model user survey and the MoPs in EMMa for the Model Users - **Scientific Researcher and Model Developer***

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
1	While using the DTA, how important is the presence of the feature - vertical queuing - ability to adhere to strict capacity constraint and hence flow never exceeds capacity, valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model Developer**	*Flow Metering - strict capacity constraint,* *Propagation - Link flows,* *Propagation - Effect of time variability in Fundamental Diagram*
2	While using the DTA, how important is the presence of the feature - horizontal queuing - ability to adhere to strict storage constraint and hence spillback into upstream links may occur, valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Traffic Spillback - strict storage constraint,* *Propagation - Queuing*
3	While using the DTA, how important is the presence of the feature - capacity drop, valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Capacity drop*
4	While using the DTA, how important is the presence of the feature - second-order effects of traffic states such as gradual increase in speeds, flows et,c valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Link-level dynamic distribution of vehicle speeds- Curve roughness factor,* *Link-level dynamic distribution of traffic flows- Curve roughness factor*
5	While using the DTA, how important is the presence of the feature - the variability in route options by means of generating routes through stochastic methods such as Monte Carlo simulations as opposed to a set of pre-defined routes provided as input by the model user, valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Presence of variable route set*
6	While using the DTA, how important is the presence of the feature - stop & go waves, valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Modeling of stop and go waves*

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
7	While focusing on the convergence of the Dynamic Assignment Module, how important is the value of the dynamic relative duality gap, valid for the respective planning horizons? Under the assumption that the model converges to a value below the pre-specified equilibrium threshold, therefore a smaller value of the duality gap would mean a better convergence. *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Dynamic Relative duality gap*
8	Looking at aspects of multi-class applicability, how important is it to include different input parameters for multiple travel modes - road-bound users such as private cars, public transport, bikes, walking etc. and trips purposes - leisure, commute etc.valid for the respective planning horizons? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Difference in Network Supply based on Modes,* *Difference in Input parameters based on different trip purposes*
9	Looking at the Dynamic Network Loading Module, what is the importance of a theoretically sound node model (merge and diverge nodes) while comparing the expected theoretical results vs actual simulated results? (Under the assumption that the node model follows the requirements stated in (Tampère, Corthout, Cattrysse, & Immers, 2011) *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Node model-merge behavior,* *Node model-diverge behavior*
10	Looking at the Dynamic Network Loading Module, what is the importance of a consistent propagation behavior in a signalized Intersection while comparing the expected theoretical results vs actual simulated results? Signalized Intersection behavior in the events of different saturation conditions, spillback conditions, merge-diverge flows etc.are evaluated in this case *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Signalized Intersection*

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
11	Looking at the Dynamic Assignment Module, what is the importance of route choice behavior with route cost variations influenced by link cost feedback in Dynamic Network Loading Module while comparing the expected theoretical results vs actual simulated results? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Route choice (general)*
12	Looking at the Dynamic Assignment Module, what is the importance of evaluating the effect of route overlap in route choice behavior by analyzing the relative differences during a MNL and PCL run of the DTA while comparing the expected theoretical results vs actual simulated results? *(On a scale from 1 - least important to 10 - most important)*	**Scientific Researcher and Model developer**	*Route choice (route overlap)*
13	While using the DTA, what is the importance of observing behavioral differences in Urban and Non-urban links, comparing the expected theoretical results vs actual simulated results? The expectation here would be that urban links have a shorter length with smaller average speed in comparison to motorway links, as people in city conditions tend to drive a lot more aggressively, resulting in bigger fluctuations in speed, density and flow in a shorter period of time. (On a scale from 1 - least important to 10 - most important)	**Scientific Researcher and Model developer**	*Fluctuation of traffic states over a series of urban and non-urban links*
14	During the model run, how important is the computational efficiency, measured in terms of run time and peak memory usage, for a specific case scenario? (On a scale from 1 - least important to 10 - most important)	**Scientific Researcher and Model developer**	*Run Time in Sec,* *Peak memory Usage in MB's*

Sl No	Question Description	Target Model User	Link to MoPs Evaluated
15	How important is the usability of the model, defined as the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component (Definition quoted from IEEE Std.610.12-1990, referred from (Seffah, Donyaee, Kline, & Padda, 2006) ? (On a scale from 1 - least important to 10 - most important)	**Scientific Researcher and Model developer**	*Familiarity,* *Simplicity,* *Navigability,* *Controllability,* *Readability,* *User guidance,* *Flexibility*



Q3.2. While using the transport model for the applications in the options below, how important would it be to incorporate the following real-world effect: blockage of vehicles in a road section as a result of traffic jams and induced congestion in preceding road sections upstream from the blockage?

(On a scale from 1 - least important to 10 - most important)

	1	2	3	4	5	6	7	8	9	10
Strategic Planning	○	○	○	○	○	○	○	○	○	○
Tactical Planning	○	○	○	○	○	○	○	○	○	○
Operational Planning	○	○	○	○	○	○	○	○	○	○

Figure 173: Sample question with response matrix used for the survey questionnaire