A Scaling-up Method for Assessing the Impacts of ITS on Traffic Efficiency

Xiaofang Huang
A Scaling-up Method for Assessing the Impacts of ITS on Traffic Efficiency

by

Xiaofang Huang

Master of Science
in Transport, Infrastructure and Logistics

at the Delft University of Technology

Student number: 4327527
Project duration: February, 2015 – August, 2015

Graduation committee:
Prof. dr. ir. S. P. Hoogendoorn (Chairman), TU Delft
Dr. ir. A. J. Pel, TU Delft
Dr. M. van Noort, TNO
Dr. J. A. Annema, TU Delft
This research is the final proof of competence for obtaining the Master of Science (MSc) degree in Civil Engineering, with a specialization in Transport, Infrastructure and Logistics (TIL) from University of Technology Delft, The Netherlands. The research is performed at TNO during the period of February 2015 until August 2015 under the supervision of the graduation committee.

The purpose of this research is to develop a scaling-up method for ITS that make direct network-wide influences to assess their large-scale impacts on traffic efficiency, and to evaluate the new method via applying it to a specific ITS application. To achieve this research objective, the research reviews the current scaling-up methods, designs a new scaling-up method and applies the new method to a specific ITS. It is concluded that the designed new scaling-up method in this study is able to assess the large-scale impacts on traffic efficiency for ITS that make direct network-wide influences.

Here, I would like to take this opportunity to express my gratitude to my supervisors, Prof. dr. ir. Serge Hoogendoorn, Dr. ir. Adam Pel, Dr. Martijn van Noort and Dr. Jan Anne Annema, for their time, patience and guidance. Prof. dr. ir. Serge Hoogendoorn, as the chairman of my thesis committee, proposed many helpful advices from a quite professional perspective in every single meeting we had. Dr. ir. Adam Pel, my daily supervisor in the university, not only provided strong supports for my thesis but also gave me useful suggestions about how to do a good research. In TNO, Dr. Martijn van Noort helped me a lot and offered useful advices. In the political research part, Dr. Jan Anne Annema provided me a lot help. I would also like to acknowledge and thank other researchers that have helped me in this study: Yusen Chen, Eline Jonkers, Taoufik Bakri, Michiel Minderhoud and Maaike Snelder. A final thanks is contributed to my dear family and my best friend Claire who offered me strong spiritual support.
Human beings today have to face a series of problems brought by transport development — severe urban congestion, increasing number of injuries and fatalities as well as global warming caused by excessive emissions. Intelligent Transport Systems (ITS), as effective tools to solve these problems, thus have drawn much attention. In the future, it is expected that more and more ITS would be developed and applied in real practice.

Before adopting ITS measures, it is necessary for policy makers to know the impacts of the ITS measure on a large scale (e.g. national/European level). In many cases, the impacts of ITS are evaluated on a much smaller scale, for example from a microscopic traffic simulation or a field experiment. These effects need to be scaled up to the larger scale. There are two known scaling-up methods. The *modelling method* can accurately represent the large scale scenario, but requires considerable effort and a large amount of data which may not be available. Furthermore, it requires a macroscopic model of the ITS, which may be a challenge to derive. The *statistical method* describes the local scenarios via situational variables (like road types, vehicle types and traffic situations), classifies the local scenarios into categories and calculates the impacts on large scale as the weighted average of the local impacts. This method is easier and faster than the modelling method. However, the statistical method is only applicable for cases which only consider categorical situational variables, because the *classification of the local scenarios into categories* is not feasible when numerical situational variables are used. As a result, the statistical method is only suitable for ITS whose impacts are on the microscopic mechanisms (e.g. speed and headway) and thus mainly affected by categorical situational variables (e.g. road type and vehicle type). A scaling-up method to assess the impacts of ITS on traffic efficiency which is generally suitable for all ITS is still missing.

To start filling this gap, this study develops a new scaling-up method for ITS that have direct network-wide influences to assess their large-scale impacts on traffic efficiency. The framework of the new scaling-up method is shown in Figure 2.2 and the graphical and mathematical interpretations are presented in Figure 2.3 and Figure 2.4. In brief, the new scaling-up method firstly chooses the suitable indicator of the impacts and situational variables, then collects needed data and builds deterministic relationships between the indicator and the numerical situational variables, at last uses scaling sideways to calculate all local impacts and aggregates the local impacts to large scale. From the theoretical perspective, the designed method is considered to be able to evaluate the impacts of ITS measures with direct network-wide influence on traffic efficiency in a large-scale scenario.

To provide an evidence of the quality of the new scaling-up method, this study applies it to a specific ITS measure, that is the on-trip dynamic navigation system. Although the final large-scale impacts of the on-trip dynamic navigation system is not calculated due to the limitation of data source, it is proved that the new method is able to accurately assess
the large-scale impacts of the on-trip dynamic navigation system with enough available data. Other findings from the case study are also valuable. For example, the choice of the indicator and the situational variables, and the built deterministic relationship can be directly adopted in other projects that study the impacts of the on-trip dynamic navigation system, which indicates the practical contribution of this study.

From a methodological perspective, the new scaling-up method is a great improvement of the current scaling-up approaches. The new scaling-up method expands the application area of scaling-up methods to ITS that have network-wide influences. Compared to the current methods, the new scaling-up method also improves the accuracy of scaling up and leads to more reliable assessments. Apart from the merits, there are also some disadvantages of the new scaling-up method, such as the possibility of more time cost and data needed, as well as the possible difficulty to explain the deterministic relationships in a sensible way.

The new scaling-up method is regarded to be with significant political value. The outputs can provide useful information to support policy making. On one hand, according to the political economy model designed by Beuthe (Figure 1.1), the impacts of ITS play an important role in making policy decisions. The impacts of ITS can directly reflect the perceived effectiveness and the perceived distribution of benefits and costs. On the other hand, based on the outputs of the new scaling-up method, there are also other policy advices that could be made. For instance, the built deterministic relationship(s) can suggest the to-be-set value of the related parameters when adopting a certain ITS measure.

For future researches, the attention could be focusing on applying the new method to more ITS measures and investigating the applicability of the new method on assessing the impacts on safety and environment. Specifically regarding the study of the on-trip dynamic navigation system, if the needed data is available, it would be beneficial to conduct a complete assessment of the large-scale impacts in a specific scenario in the future. In addition, the influences of other situational variables besides the considered situational variables could also be taken into account. Furthermore, a more specific classification of network structure is expected in future researches.
CONTENTS

List of Figures ix

List of Tables xiii

1 Introduction 1
   1.1 Background 1
   1.2 Current scaling-up methods 5
   1.3 Problem definition 9
   1.4 Research objective 11
   1.5 Research question 11
   1.6 Research approach and readers’ guide 12

2 The new scaling-up method 13
   2.1 Design of the new scaling-up method 13
   2.2 Instruction for using the new scaling-up method 23
   2.3 Conclusion 25

3 Applying the new scaling-up method to a specific ITS application 27
   3.1 Before applying the new scaling-up method 27
   3.2 Determination of the indicator of the impacts on traffic efficiency 29
   3.3 Choice of the situational variables 36
   3.4 Data collection 42
   3.5 Building the deterministic relationship 61
   3.6 Scaling up procedure for the on-trip dynamic navigation system 89
   3.7 Conclusion 90

vii
**Contents**

4 Evaluation of the new scaling-up method 93
  4.1 Evaluation from a methodological perspective 93
  4.2 Evaluation from a political perspective (political contribution) 96
  4.3 Conclusion 97

5 Conclusions and recommendations 99
  5.1 Conclusions 99
  5.2 Recommendations 102

Bibliography 105

A Literature related 109

B DYNASMART-P FEATURES 113
  B.1 Demand Representation 113
  B.2 Traffic Control 113
  B.3 Different Vehicle Types 113
  B.4 Multiple User Classes 114
  B.5 Operational modes 114

C Maps of networks 119
# List of Figures

1.1 A political economy model of transport innovations [Beuthe, 2004] ........................................... 4
1.2 An example of categories in the direct method (euroFOT) .......................................................... 6
1.3 The conceptual model of SCENIC [Malone et al., 2014] ................................................................. 7
1.4 Scaling-up approach used in Amitran [Mans et al., 2013] ............................................................. 8
1.5 An example of scaling up using statistics (Amitran) [Jonkers et al., 2014] ................................. 9
1.6 Research approach and readers’ guide .......................................................................................... 12

2.1 An example of wrong use of the relationship defined in the local scenarios for the large scenario ................................................................................................................... 18
2.2 The framework of the new scaling-up method ............................................................................ 19
2.3 The graphic interpretation of the new scaling-up method .......................................................... 21
2.4 The mathematical interpretation of the new scaling-up method ............................................. 22

3.1 Influence of highways in Network 1 (Pilot 1-1) ......................................................................... 52
3.2 Influence of actuated signals in Network 1 (Pilot 1-2) ............................................................. 53
3.3 Influence of the penetration rate on the effectiveness of the on-trip dynamic navigation system (Pilot 2) .................................................................................................................... 54
3.4 Influence of the congestion level on the effectiveness of the on-trip dynamic navigation system (Pilot 3) .................................................................................................................... 55
3.5 Influence of information update interval on the effectiveness of the on-trip dynamic navigation system (Pilot 4) .................................................................................................................... 56
3.6 Influence of the network density on the effectiveness of the on-trip dynamic navigation system (Pilot 5) .................................................................................................................... 58
3.7 Examples of fitting data by quadratic functions ........................................................................ 63
3.8 Examples: residual analysis of fits using Function 3.7 ............................................................ 65
3.9 Range of value of R-squared under different congestion levels (XA2 network, 2-minutes information update interval) ................................................................. 66
List of Figures

3.10 An Example of fitting data by Function 3.7 (FW2 network, 2-minutes information update interval) .................................................. 67
3.11 Value of $a_1$ under different congestion levels ........................................... 68
3.12 Value of $a_2$ under different congestion levels ........................................... 69
3.13 An example of fitting data by Function 3.9 (CD3 network, 2-minutes update interval) ................................................................. 70
3.14 An Example: residual analysis of fits using Function 3.9 (CD3 network, 2-minutes update interval interval) ........................................... 71
3.15 Examples of residual analysis (investigating the influence of the information update interval) .................................................. 73
3.16 Residuals analysis of fits using Function 3.12 (the horizontal axis shows the same case number as Figure 3.15) ........................................... 75
3.17 Residuals analysis of fits using Function 3.12 ........................................... 76
3.18 Residuals analysis of fits using Function 3.12 (all networks) ................... 79
3.19 Residuals when fitting data by Function 3.9 (all networks) ................. 80
3.20 Proportion of absolute residuals in each range when fitting data by Function 3.12 (all networks, all information update intervals) ........................................... 82
3.21 Proportion of absolute residuals in each range when fitting data by Function 3.12 (all networks) ........................................... 82
3.22 Proportion of absolute residuals in each range when fitting data by Function 3.19 (all networks) ........................................... 83
3.23 Evaluation of applying values of parameters in Function 3.12 from a network to another ................................................................. 85
3.24 Evaluation of applying parameters in Function 3.9 from a network to another 87
3.25 Scaling up procedure when different information update intervals are applied 89
3.26 Scaling up procedure when a fixed information update interval is applied 89

A.2 Congestion survey results [Bertini, 2006] ........................................... 110
A.3 Traffic congestion measures and their suitability to assessment criteria [Aftabuzaman, 2007] ................................................................. 110
A.5 Common Congestion Indicators [Litman, 2013] ........................................ 111
B.1 DYNASMART-P (One-step Simulation Assignment) procedure ....... 115

C.1 Map of Network 1 ........................................ 120
C.2 Map of Network 2 (FW2 network) .......................... 121
C.3 Map of Network 3 (FW1 network) .......................... 122
C.4 Map of Network 4 (FW3 network) .......................... 123
C.5 Map of XA1 network ........................................ 124
C.6 Map of XA2 network ........................................ 125
C.7 Map of XA3 network ........................................ 126
C.8 Map of CD1 network ........................................ 127
C.9 Map of CD2 network ........................................ 128
C.10 Map of CD3 network ....................................... 129
C.11 Map of AM1 network ....................................... 130
C.12 Map of AM2 network ....................................... 131
LIST OF TABLES

3.1 Different ITS and their main /first-order impact level(s) ............................ 28
3.2 Assessment of the measures of traffic efficiency based on the criteria ............ 35
3.3 An example of the inability of the change of TTI to indicate the impacts of the on-trip dynamic navigation system .......................................................... 36
3.4 Assessment of situational variables based on the criteria ............................ 40
3.5 Influence of random seeds in Network 2 (Pilot 1-3) ................................. 53
3.6 The range of values of situational variables in Pilot 2 ............................... 54
3.7 The range of values of situational variables in Pilot 3 ............................... 55
3.8 The range of values of situational variables in Pilot 4 ............................... 56
3.9 The range of values of situational variables in Pilot 5 ............................... 57
3.10 The information of networks analysed in the study ..................................... 60
3.11 The range of values of situational variables (except the network density) in the study .......................................................... 60
3.12 Values of R-squared (fitting with Function 3.7) ..................................... 64
3.13 Values of R-squared (fitting with Function 3.9) ..................................... 70
3.14 Values of R-squared (investigating the influence of the information update interval) .......................................................... 72
3.15 Values of R-squared (fitting with Function 3.12) ..................................... 74
3.16 MAE (fitting with Function 3.12) .......................................................... 74
3.17 Values of R-squared (fitting with Function 3.9 for homologous networks) .... 78
3.18 Values of R-squared (fitting with Function 3.9 for all networks) ............... 79
3.19 MAE (fitting with Function 3.19 for all networks) ..................................... 83
This chapter is an introduction for the study. In the first place, the background of the study is explained in Section 1.1, stating the development of Intelligent Transport Systems (ITS) and the necessity of using scaling-up methods. Afterwards, the existing scaling-up approaches are introduced in Section 1.2, and through analysing these methods, the current problem (gap) is defined in Section 1.3. The research objective and research questions are thus proposed in Section 1.4 and Section 1.5. At last in Section 1.6, the research approach is presented and reader’s guide is shown.

1.1. BACKGROUND

As a promising solution to reduce the negative effects bought by transport development, ITS have been developed rapidly in the past few decades around the world, so it is expected that more and more ITS will be developed and applied in practice in the future (Section 1.1.1). Before adopting a new ITS measure, it is important for policy makers to know the impacts of the ITS application on the large scale (Section 1.1.2). Nonetheless, in many cases, the impacts of ITS are obtained on a much smaller scale, so scaling-up methods are needed to scale up these effects to the large scale (Section 1.1.3).

1.1.1. ITS: A PROMISING SOLUTION

Transport plays a very important role in the course of human development. In the past centuries, from horses to self-driving cars, transport has been developed rapidly with the development of technology. One fact is the continuously growing number of car owner-
ships. According to Ward’s research, which looked at government-reported registrations and historical vehicle-population trends, global registrations had reached 1,015 billion by 2010 [Sousanis, 2011]. But human beings today have to face a series of problems brought by transport development — severe urban congestions, increasing number of injuries and fatalities as well as global warming. ITS, which are effective to reduce the negative consequences of transport development, have thus drawn a lot attention throughout the world over the past few decades.

In the United States, each state owns an ITS chapter which holds a conference yearly to promote ITS-related technologies and ideas. In addition, several national organizations, for example ITS America, are dedicated to developing ITS for improving the nation's surface transportation system. In Japan, because of the massive investment from the government, many kinds of ITS have been successfully adopted around the country, which benefits the public greatly and enables Japan to play a crucial role in ITS field [An et al., 2011]. In Europe, as early as in 1985, European Road Transport Telemetric Implementation Coordination Organization (EUREKA) had been established in order to promote the cooperation between government and private organizations in the research and development of ITS [An et al., 2011]. In the Netherlands, by 2010, a total of 2,587.35 kilometres of national roads had been equipped with standard functionality (ITS), and additional 1,278.50 kilometres with monitoring equipment only. This is approximately 68% of the total road length [Connekt, 2011]. Apart from that, up to 2010, 135 roadside DRIPs (Dynamic Route Information Panels) and 104 Ramp meter signals have been installed [Connekt, 2011] in the country.

The development of ITS is also directly supported by the policies. For example, in Europe a legal framework Directive 2010/40/EU was adopted on 7 July 2010 to accelerate the deployment of ITS. This directive aims to establish interoperable and seamless ITS services while leaving Member States the freedom to decide which systems to invest on [European Commission, 2014]. Four priority areas have been determined in this directive: i) optimal use of road, traffic and travel data; ii) continuity of traffic and freight management ITS services; iii) ITS road safety and security applications; and iv) linking the vehicle with the transport infrastructure.

According to the information above, it is not hard to see that government makes lots of efforts to develop ITS around the world. Predictably, in the future, more and more ITS will be developed and applied in practice with the encouragement of policies.

1.1.2. Importance of Assessing the Impacts of ITS on the Large Scale

As concluded in the previous section, there would be more and more ITS proposed in the future. However, before applying a new ITS measure, it is important for policy makers to
know the impacts of ITS.

The budget that can be invested to develop ITS measures is often limited in a certain area, so only some of the proposed ITS measures could be practically adopted. Even if there are several ITS measures can be developed within the budget, policy makers still need to carefully select the ITS measures that would be adopted in order to maximize the benefits for the public. As a result, policy makers always need and want to know if the ITS measure is deserved to be adopted in order to make a decision. One widely used tool that can help policy makers to make the decision is the political economy framework developed by Beuthe [Beuthe, 2004], shown in Figure 1.1. Beuthe insists that a transport innovation, for instance an ITS measure, can be adopted only if it is technical, socially and politically feasible.

If a transport innovation is technically feasible, it should be able to work from a technical perspective, in other words, it can meet all technical requirements. An innovation could be considered socially feasible if a major of voters would like to support it, which indicates that there is a demand for the innovation. An innovation which is supported by the policy makers could be considered politically feasible. The political feasibility of an innovation partially depends on its social feasibility, indicating that the policy makers do take account the public preference.

It is not hard to find that the impacts of transport innovations (e.g. ITS measures) play a crucial role in the model developed by Beuthe, because both of the perceived effectiveness and the perceived distribution of benefits and costs can be significantly influenced by the impacts of transport innovations.

The perceived effectiveness of an innovation is related to the degree of the innovation reaching the policy goal(s) (e.g. less congestion, few injuries, and higher air quality). The better the innovation achieves the political goal(s), the higher the perceived effectiveness is. So the perceived effectiveness is directly reflected by the effectiveness of the innovation (i.e. the impact of the innovation).

The perceived distribution of benefits and costs can be explained from two perspectives. This term on one hand can represent the ratio of costs to benefits, on the other hand, can express the distribution of costs (benefits) among different groups. The impacts of the innovation, as basic parts of the benefits, are influential to the ratio of costs to benefits. Additionally, if the impacts of the innovation are studied among different groups, they can also affect the distribution of benefits. So the impacts of the innovation is an influential factor for the perceived distribution of benefits and costs of the innovation.

The model built by Beuthe indicates that the impacts of a transport innovation are quite important for making the political decision that if this transport innovation should
be adopted. Therefore, before adopting ITS measures, policy makers would like to know their impacts. Furthermore, considering the political decision of adopting ITS measures is often made on a large scale (e.g. country level and EU level), the impacts of ITS on the large scale are what the policy makers want and need to know.

1.1.3. **Need for scaling-up methods**

Policy makers would like to know the impacts of ITS on the large scale (e.g. country level and EU level), however, in many cases, the impacts of ITS are obtained on a much smaller scale (i.e. local scale such as road level and city level) from microscopic traffic simulations or field experiments. These effects need to be scaled up to the large scale. Currently, in some studies, researchers tries to use the **scaling-up method** to assess the impacts of ITS on the large scale. The **scaling-up method** is a procedure of translating the impacts of ITS on the local scale to the impacts on the large scale by scaling up.

The definitions of the terms of ‘the local scale’ and ‘the large scale’ are not fixed. These two terms are relative concepts: ‘the local scale’ is the lower level and ‘the large scale’ is
the higher level. For example, if the large scale is defined with EU level, the local scale is possibly the level of European countries. But if the large scale represents the level of the Netherlands, the local scale could be the lower level (e.g. level of Dutch cities). The definitions of the two scales depends on the view of the projects or researches. Specifically, in this study, the following definitions are made to clarify the expression:

- **the large scale**: the larger scale compared to the local scale, often indicating country level or EU level.
- **the large scenario**: the corresponding scenario on the large scale. Generally there should be only one large scenario in a study as the target scenario, for example a country or Europe.
- **the large-scale impacts**: the impacts of ITS on the large scale (the impacts of ITS in the large scenario).
- **the local scale**: the relative lower level compared to the large scale.
- **the local scenarios**: the scenarios on the local scale, included in the large scenario.
- **the local impacts**: the impacts of ITS on the local scale (the impacts of ITS in the local scenarios)

The scaling-up method is considered as a desirable approach to assess the impacts of ITS on the large scale. It takes full advantage of the outcomes in microscopic traffic simulations or field experiments. Besides, there is no need to build and run a complex macroscopic simulation model, thus less time is required for collecting data and verifying model.

### 1.2. Current scaling-up methods

Although the scaling-up method is a promising approach to assess the impacts of ITS on the large scale, there are few records of projects using this type of method. The first reason is that the scaling-up method is a new emerging approach. It comes to the writer’s knowledge that the first project in Europe using the scaling-up method to assess the impacts of ITS is *euroFOT* which was started in 2008 [euroFOT, 2015b]. The limited number of projects is also a result of the limited budget. There is a requirement for abundant funding to complete a big project which studies the impacts of ITS on country/EU level. So generally, projects are funded by government or more often by European organizations like European Commission. But the financial supports from government or organizations are limited, so only handful projects could be proceeded.
Even though the number of projects is small, the scaling-up approaches used in these projects are still of great referential value. In this section, three projects using scaling-up to study the impacts of ITS at EU level are discussed.

1.2.1. Project - euroFOT (started in 2008)

The project euroFOT established a comprehensive, technical, and socio-economic assessment programme for evaluating the impacts of ITS on safety, environment, and driver efficiency [ euroFOT, 2015b ]. In this project, only intelligent vehicle innovations, like Adaptive Cruise Control and Curve Speed Warning, were studied [ euroFOT, 2015a ].

There are two scaling-up methods proposed in euroFOT project. One is called the **modeling method** which uses macroscopic simulation models and another is called the **direct method** which uses statistics. In the modelling method, the outputs of the microscopic simulation are the inputs for the macroscopic simulation, and the scaling-up results could be computed by the macroscopic simulation model. However, sometimes a macroscopic model is not available, the direct method thus can be used as a supplement. In the direct method, the impacts on the local scale are collected by simulations and split for the situational variables. For example, from the simulation, the travel times for two scenarios in six categories (as shown in Figure 1.2) and the number of kilometres driven in each category for each scenario can be known. The data of driven kilometres for each category is also available for the whole Europe. A weighted travel times on EU level for each scenario could be calculated, and a comparison between the two scenarios can be made. (The information in this paragraph is based on [Faber et al., 2011].)

![Figure 1.2: An example of categories in the direct method (euroFOT)](image)

In the project, the direct method was more favoured by the researchers — "in euroFOT a direct method for scaling up will be used" [Faber et al., 2011]. Nonetheless, when the direct method was applied in the project, the number of chosen categories was so limited
that the considered situations was far fewer than the actual possible situations. Only three road types, two congestion levels and two penetration rates were taken into account.

1.2.2. **PROJECT - DRIVE C2X (STARTED IN 2011)**

DRIVE C2X, an European Integrated Project, developed a testing scaling-up methodology and assessed the impact of cooperative systems on users, environment and society [FOT-Net, 2015]. SCENIC, a standard tool developed by TNO, was used to scale up the impacts of cooperative systems in DRIVE C2X and its conceptual model is shown in Figure 1.3.

SCENIC assumes that the impacts of an ITS application on the local scale, which is defined as *local impacts*, are known. A local scenario describes a situation and the circumstances in the local scenario are characterized by *situational variables*. As output, the tool provides the impacts of this ITS for a target scenario which are called *target impacts*. Both of the local impacts and the target impacts are defined as changes in the *performance indicator* and the same performance indicator is used for input and output. In order to translate the local impacts to the target impacts, external data are needed to weight the local scenarios in the target scenario. The impacts of this ITS application in the target scenario is thus calculated as the weighted average of the impacts of the ITS application in the local scenarios. (The information in this paragraph is based on [Malone et al., 2014].)

![Figure 1.3: The conceptual model of SCENIC](Malone et al., 2014)

It is not hard to find that the SCENIC is quite similar to the direct method mentioned in euroFOT. This limitation of SCENIC is also similar to that of the direct method. Although no restrictions are set on situational variables in SCENIC, according to the conducted cases in DRIVE C2X project, the number of chosen scenario parameters is still quite small — the local scenarios were generated in terms of only three types of roads (highspeed roads, rural roads and urban roads) and two demand levels (low demand and high demand).
1.2.3. Project - Amitran (Started in 2011)

Amitran, a project led by TNO and co-funded by European Commission, defined a reference methodology to assess the impacts of ITS on CO₂ emissions [TNO, 2014].

Two scaling-up methods are proposed in Amitran (also shown in Figure 1.4):

- Scaling up using a (macroscopic) multi-modal traffic simulation model on large scale (e.g. European or country level)

- Scaling up using statistics (direct method)

Figure 1.4: Scaling-up approach used in Amitran [Mans et al., 2013]

When scaling up by using a macroscopic model, if the local effects are already known, the impacts are performed directly with a macroscopic traffic simulation model, otherwise, the local effects need to be calculated with a microscopic simulation model firstly. After that, a second run is performed with the macroscopic model to account for the second order effect. When second effects, like influence on demands, are considered important, scaling up using a macroscopic model is a good method because it can take into account specific circumstantial differences, but this method can be used only when the large-scale model is available [Mans et al., 2013].

When scaling up using statistics, the local impacts are collected for instance from experiments, literature or microscopic simulations and distinguished by different situational
variables firstly, for example, 10% reduction of CO₂ emission (impacts) on highways (situational variable). After that, based on statistics data like kilometres driven for each type of roads, the impacts on the large scale could be assessed. An example is given below. Note that to use this method, statistics need to be available for all local scenarios [Mans et al., 2013].

The two scaling-up methods applied in Amitran are quite similar to those designed in euroFOT, but the scaling-up methods in Amitran were designed specifically for assessing the impacts of ITS on CO₂ emission and the authors did not mention its applicability for assessing the impacts of ITS on safety and traffic efficiency.

1.3. PROBLEM DEFINITION

According to Section 1.2, there are two main scaling-up approaches in current studies. One can be called the **modelling method** — using the local impacts as the inputs for the large-scale model then calculating the impacts of ITS on the large scale. The other one can be called the **direct method** — describing the local scenarios via situational variables (e.g. road types, vehicle types and traffic situations), classifying the local scenarios into categories and calculating the large-scale impacts as the weighted average of the local impacts.

The modelling method is more traditional compared to the statistics method and its results are expected to be closer to the reality. However, to use this method, the impacts of ITS in all local scenarios need to be known, which may not be available. Furthermore, it requires a large-scale model, which may be a challenge to derive. The modelling method is not widely used in the previous projects.

The direct method is frequently used the cited three projects. However, it is noted that
in all the three projects, when applying this method, only three types of situational variables (road type, vehicle type and traffic situations) are taken into account. Only using these three situational variables might be enough in some cases, but in other cases, it is not acceptable.

When evaluating the impacts of ITS on safety or environment, such a simple set of situational variables might be satisfied because vehicle type and road type are the most important two influential factors in most cases. For example, ITS often affect $CO_2$ emission by changing the driving dynamics like the acceleration and deceleration process. Considering the driving dynamics are closely related to speed limits and the vehicle characteristics, road type and vehicle type are the most influential factors for the impacts of ITS on environment.

When assessing the impacts of ITS on traffic efficiency, using such a simple set of situational variables could also be fine if the impacts of ITS are microscopic (i.e. the impacts are on the microscopic mechanisms like the lane choice, the speed, and the headway). Because the microscopic effects are often road-related and vehicle-related instead of network-related, using road type and vehicle type as the situational variables is generally enough. However, if the impacts of ITS are directly on the macroscopic mechanisms in the network such as mode or route change (i.e. the ITS have direct network-wide effects), only considering these three situational variables is far from enough. For instance, if an ITS application can affect the route choice, then its impacts on traffic efficiency would mainly rely on the number of alternative routes and the differences in travel times among alternatives but not on the road type or the vehicle type. In such a case, more situational variables other than road type, vehicle type and traffic situations need to be selected.

More types of situational variables should be considered when assessing the impacts of ITS on traffic efficiency. Nonetheless, the direct method might not be applicable any more when other situational variables are selected. For example, if the number of alternative routes is one of the situational variables, it seems not possible to classify the local scenarios into categories because the number of alternative routes is a value instead of a type, thus the direct method cannot be used in such a case.

According to the limitation of the modelling method and the direct method, the current problem can be defined as follows:

There is a need for scaling-up methods to assess the impacts of ITS on the large scale, but when assessing the impacts on traffic efficiency, the current scaling-up approaches are only applicable for ITS whose impacts are on microscopic mechanisms. A scaling-up method which is suitable for all types of ITS to assess the impacts on traffic efficiency is still missing.
1.4. **Research Objective**

Based on the previous section, the main flaw of the current scaling-up methods is that they cannot be properly applied to ITS which have influences directly on network level. So, if a new scaling-up method which is able to evaluate the impacts of ITS that have direct network-wide effect on traffic efficiency could be designed, this gap would be filled to some extent. Additionally, after designing of the new scaling-up method, it would be favoured to apply the method to a specific case in order to provide evidence of the quality of the new method. Therefore, the objective of this study is:

*To develop a scaling-up method for ITS that make direct network-wide influences to assess their large-scale impacts on traffic efficiency, and to evaluate the new method via applying it to a specific ITS application.*

Note that **ITS that make direct network-wide influences** are the ITS whose direct impacts are on the macroscopic mechanisms in the network like demand generation, modal split, departure time choice and route choice. For example, traffic information system is one of such ITS.

1.5. **Research Question**

To achieve the research objective, the following research questions (RQ) need to be answered:

- **RQ1: What are the current scaling-up methods?**
  - Provide good references for designing a new scaling-up method.

- **RQ2: What is the new scaling-up method?**
  - Result in the new scaling-up method.

- **RQ3: Is the new scaling-up method applicable for a specific ITS measure?**
  - Provide evidence of the quality of the new method;
  - Develop a detailed scaling-up approach for a specific case;

- **RQ4: How is the quality of the new method from a methodological perspective?**
  - Evaluate the new scaling-up method from a methodological perspective

- **RQ5: What is the contribution of the new method to policy making?**
  - State the value of the new method from a political perspective
1.6. **RESEARCH APPROACH AND READERS’ GUIDE**

The research approach for the whole study is thus developed according to the five research questions, shown in Figure 1.6.

![Figure 1.6: Research approach and readers’ guide](image)

In **Chapter 1**, the current scaling-up methods are introduced and discussed briefly, which results in the current research problem (gap). With the goal of filling the current gap, the new scaling-up method especially for ITS with direct network-level influences is developed in **Chapter 2**. Then in **Chapter 3**, in order to primarily investigate the quality of the designed new scaling-up method, the new method is applied to a specific ITS. According to the specific application of the new scaling-up method, evaluations of the new scaling-up method is made from a methodological perspective (the merits and flaws) as well as from a political perspective (the main contributions to political making) in **Chapter 4**. At last in **Chapter 5**, conclusions and recommendations for the whole study are drawn.
This chapter explains how the new scaling-up method is developed and introduces the new scaling-up method. In section 2.1, it discusses that how the new method is proposed and shows the framework as well as the mathematical and graphical interpretation of the new scaling-up method. In section 2.2, for better applying the new method, some advices (examples) for how to proceed each step are discussed. In the last section, the conclusions for this chapter are drawn.

2.1. Design of the New Scaling-up Method

To design a new method, it is generally needed to know the requirements which should be met. Besides, learning from strong points and offsetting weakness of the current methods is often considered as a useful way to design a new method. With the goal of meeting requirements and filling the current gap, a new method could be proposed.

Therefore, in the designing of the new scaling-up method, the basic idea is: (1) defining the required functionality that a qualified method should have, that is, the requirements towards the new method; (2) analysing the current scaling-up methods in order to see if they have the required functions and how they realize them; and (3) arguing what improvements the new method can make and by what steps it can achieve the improvements.
2.1.1. Requirements that should be met by the new scaling-up method

The fundamental concept of scaling-up is scaling up the impacts of ITS on local scale to the impacts on large scale, so it is not hard to infer that the basic requirement for any scaling-up methods is the ability to scale up the local impacts to the large-scale impacts.

In addition, according to the research objective, the new method is designed to assess the impacts of ITS that have direct network-wide influences on traffic efficiency, so the new method should be able to applied to the target ITS type.

Therefore, two main requirements that the new scaling-up method should meet are:

- **Requirement 1**: The new scaling-up method should be able to scale up the local impacts to the large-scale impacts.

- **Requirement 2**: The new scaling-up method is applicable for ITS that have direct network-wide influences on traffic efficiency.

2.1.2. To meet Requirement 1: Scaling up local impacts to large-scale impacts

In this section, it is analysed that whether the current scaling-up methods can meet the defined Requirement 1 and how they realize them. The flaws of the current methods regarding the requirement are thus discussed, which indicating what improvements the new method can make. To achieve the improvements, the steps that should be included in the new scaling-up method are proposed.

2.1.2.1. Necessity of determining a suitable indicator of the impacts

It is undoubted that the main issue should be solved by scaling-up methods is how to translate the local impacts to the large-scale impacts, but before the ‘translation’, it is also necessary to know what the local impacts are.

The local impacts seem not difficult to be known if needed data is available. However, it should be noted that ‘the impacts’ is a very general term. The impacts of ITS could be on various aspects (e.g. safety, environment and traffic efficiency). The impacts on each aspect could be represented in different ways, for example, if the impacts are on safety, they can be interpreted by ‘the reduction of fatality’, ‘the reduction of injuries’, or ‘the reduction of rear-end collisions’. Therefore, before collecting the data of the local impacts, it is necessary to determine that what measures should used to describe the impacts in other words what the
2.1. DESIGN OF THE NEW SCALING-UP METHOD

Indicators of the impacts are. Note that in a project (study), for a specific ITS application, often only one indicator of the impacts on traffic efficiency is needed, because multiple types of large-scale impacts which are represented in various indicators are generally not required.

Furthermore, the determination of the indicator of the impacts of an ITS application should be done carefully. A ‘wrong’ indicator cannot reflect the real effectiveness of the studied ITS application. For instance, if the target ITS application makes an influence mainly on queue length, using travel time as the indicator might not show the effectiveness. To make an accurate assessment, it is not only necessary to determine the indicator but also essential to choose a suitable one. The basic criterion for the suitable indicator is that a significant change in the indicator could be seen after adopting the ITS application.

However, in the current scaling-up methods, the determination of the indicator does not receive enough attention. On one hand, the considered indicators in the previous studies are very limited. For traffic efficiency, only speed, travel time and delay are mentioned, while there are lots of other indicators can interpret traffic efficiency. On the other hand, none of the current approaches consider the step of finding suitable indicators, and in the reports referring the current scaling-up methods, how the indicators are chosen is not described either.

The ignorance of the importance of determining a suitable indicator for the studied ITS application is thus a flaw of the current scaling-up methods. To fill the gap, the new method should have a step of determining a suitable indicator of the impacts of the studied ITS application on traffic efficiency.

2.1.2.2. NECESSITY OF CHOOSING SUITABLE SITUATIONAL VARIABLES

Knowing the impacts in the local scenarios, to meet Requirement 1, the issue now is how to translate these local impacts to the impacts on the large scale.

The current scaling-up methods achieve such a transformation by the following process. Firstly, researchers hold that the impacts of an ITS application in a certain scenario are also expected to be valid in comparable situations elsewhere [Mans et al., 2013]. But since each scenario is very specific to some extent, it cannot be seen at one glance if two scenarios are comparable or not. Therefore, situational variables (e.g. road type, vehicle type and traffic state), are used to describe the scenarios. For example, if the chosen situational variables are road type and vehicle type, a local scenario could be described with ‘urban road; truck’. If two local scenarios are under the same classification in every situational variable, it could be stated that these two scenarios are comparable and the impacts
in one scenario are also valid in another scenario. For example, if one ITS application is able to reduce 10% carbon emission for every truck on a highway, then it is expected that this ITS application can also reduce 10% carbon emission for every truck on another highway. The comparable scenarios could be divided into the same category. The large-scale impacts thus could be calculated as the weighted average of the local impacts in every category. By such a process, the current methods scale up the local impacts to the large scale.

Using situational variables to distinguish the scenarios is indeed a wise option. Additionally, choosing suitable situational variables are very significant. Although there are many types of situational variables could be considered, the number of situational variables that are taken into account is very limited in order to avoid massive data collection and calculation, so the choice of situational variables should be made carefully. A good choice of situational variables can not only lead to an accurate estimation of the large-scale impacts but also possibly reduce the computation time. However, no details about how to determine the situational variables are mentioned in these approaches. Only in Amitrans, it is simply stated that "the situational variables are expected to have the largest impact, the possibility of measuring the different situations, and the model capabilities" [Jonkers et al., 2014].

To offset this flaw of the current approaches, in the new scaling-up method, another step that should be included is choosing suitable situational variables.

2.1.3. To meet Requirement 2: applicability for ITS with direct network-wide influences

According to the last section, to meet Requirement 1, the new scaling-up method needs to contain at least two steps: determining a suitable indicator of the impacts of the studied ITS application and choosing suitable situational variables. Additionally, it is also found that the current methods can achieve the translation from the local impacts to the large-scale impacts by using situational variables, and the impacts on the large scale could be calculated as the weighted average of the local impacts in every category. But it is unknown that if such a translation process is also applicable for ITS that make direct network-wide influences on traffic efficiency. In this section, the analysis is made to figure out what a translation process can make the new scaling-up method to be applicable for ITS with direct network-wide influences.
2.1.3.1. **Necessity of building a deterministic relationship**

The main reason why the current scaling-up method could use the process mentioned in Section 2.1.2.2 is that only *categorical situational variables* (e.g. road type and vehicle type) are considered, since categorical situational variables make it possible to divide scenarios into categories. However, to accurately describe network-wide characteristics, categorical situational variables are not enough.

For example, if using the categorical situational variable to describe the traffic condition and simply classifying all local scenarios into two categories — congested and free-flow, under any situation, there are two possible traffic conditions. But in reality, traffic conditions in various scenarios are so distinct that reducing them into only two types would greatly eliminate the differences among scenarios.

In other words, only using categorical situational variables might make the descriptions of the network-wide characteristics not precise in certain situations. Therefore, *numerical situational variables* can be taken into account in order to precisely describe the network-wide characteristics and thus make the scaling-up method applicable for ITS with direct network-wide influences.

Nevertheless, if numerical situational variables are considered, the current scaling-up process, which needs to classify the local scenarios into categories, cannot work any more, because it is difficult even impossible to categorize numerical values. So the translation approach used in the current scaling-up methods is not applicable for ITS that have direct network-wide influences.

To make the use of numerical situational variables feasible, one possible way is *building a deterministic relationship* between the numerical situational variables and the indicator of the impacts. Based on the built deterministic relationship, the large-scale impacts could be calculated directly via the numerical situational variables.

A deterministic relationship implies that there is an exact mathematical relationship or dependence between related elements, which indicates that a deterministic relationship can only be built between the numerical situational variables and the indicator of the impacts. So building a deterministic relationship is not feasible for categorical situational variables. In other words, when both numerical and categorical situational variables are selected in one case, building only one deterministic relationship cannot relate all situational variables to the indicator of the impacts. For that matter, it might be better to *build different deterministic relationships in various categories* while the categories are decided by the categorical situational variables.
2.1.3.2. THE NECESSITY OF SCALING SIDEWAYS

Through building a deterministic relationship, the large-scale impacts are expected to be calculated directly by the relationship. However, sometimes, it could be quite risky to apply the deterministic relationship defined on the local scale directly to the large scale, especially when the generated relationship is not linear.

Figure 2.1 shows an example where the direct use of the deterministic relationship to the large scale results in a wrong estimation of the large-scale impacts. In the example, the large scenario only has two local scenarios (Scenario 1 and 2). The indicator of the impacts \(Y\) is *average delay per cap* and the numerical situational variable \(a\) is *average speed*. The relationship between the indicator and the numerical situational variable is defined as \(Y = a^2\). It is not hard to infer that the value of the situational variable \(a_3\) for the large scenario should be equal to the average of the values of the situational variable in two local scenarios \((a_1 + a_2)/2\), and the large-scale impacts \(Y_3\) should also be the average of the local impacts \((Y_1 + Y_2)/2\) which is equal to \((a_1^2 + a_2^2)/2\). But if the large-scale impacts are directly calculated by the deterministic relationship, the value of the indicator in the large scenario would be \(a_3^2 = ((a_1 + a_2)/2)^2\), which is not consistent with the actual impacts \(((a_1^2 + a_2^2)/2)\). This example in fact suggests that in some cases, applying the deterministic relationship defined in the local scenarios directly for the large scale can lead to a wrong estimation of the large-scale impacts.

![Figure 2.1: An example of wrong use of the relationship defined in the local scenarios for the large scenario](image)

If the deterministic relationship defined in the local scenarios cannot be directly applied on the large scale, it seems all local impacts need to be known to evaluate the large-scale impacts. Nonetheless, it is time consuming to collect all local impacts which are often not available. In addition, if the impacts in all local scenarios are already known, the large-scale impacts can be calculated by simply summing up, and hence designing a scaling-up method would be meaningless. Under such a situation, **scaling sideways** is taken into account in the new scaling-up method. Scaling sideways means *scaling the impacts in one
local scenario to another local scenario. Based on the defined relationship in some local scenarios, the impacts of ITS in the rest of the local scenarios could be quickly calculated by scaling sideways. Then, naturally, the impacts of ITS in the large scenario could be calculated by aggregating all local impacts.

### 2.1.4. Framework of the New Scaling-Up Method

According to the analysis above, to meet the two defined requirements in Section 2.1.1, the new method should have the following steps (at least): determining a suitable indicator, choosing suitable situational variables, building a deterministic relationship and scaling sideways. The framework of the new scaling-up method is thus built as Figure 2.2.

![Figure 2.2: The framework of the new scaling-up method](image)

The indicator and the situational variables should be determined firstly. After determining the indicator and situational variables, to build a deterministic relationship between the numerical situational variables and the indicator, it is necessary to collect the data of the values of the numerical situational variables and the impacts in some local scenarios. So the next step should be to collect data.

Specifically, if not all situational variables are numerical in other words certain situational variable(s) are categorical, the data collection should be conducted by categories.
and the following step is trying to **build a deterministic relationship (in each category)** based on the collected data (in each category). However, although the situational variables are assumed influential to the impacts, it is not necessarily possible to build a mathematical relationship between them. For example, the influence of certain situational variable might be too slight to build a sensible relation to the impacts, or the deterministic relationship might be too complex to be shown in an understandable function. If in certain category, not a sensible deterministic relationship could be found, other situational variables can also be examined. So there should be a **feedback loop** if failed to build a deterministic relationship. Certainly, if in every category, a deterministic relationship can be found, **scaling sideways (in each category)** can be proceeded to calculate the impacts in the other local scenarios whose impacts are unknown. Note that in the new scaling-up method, the criterion that a deterministic relationship can be found can be interpreted as a deterministic relationship, which can estimate the impacts through the numerical situational variables with acceptable accuracy, can be built.

On the other hand, if all situational variables are numerical, the steps to **collect data** and to **build a deterministic relationship** do not need to be done by categories. But similarly, if a deterministic relationship could not be found, there should be a **feedback loop** to re-choose suitable situational variables. If a deterministic relationship can be determined, the next step is to **scale sideways** which can calculate the unknown local impacts.

Since the unknown local impacts can be calculated in the step of scaling sideways, the impacts in all local scenarios are known after this step. So at the last, by the step of **aggregating the local impacts to the large scale**, the impacts in the large scenario can be known.

### 2.1.5. Graphical and Mathematical and G Interpretations of the New Scaling-up Method

In the previous section, the basic framework of the new scaling-up method is shown. For better understanding the concept of the new scaling-up method, a graphical description of the core idea is also made and shown in Figure 2.3 as a reference.

Furthermore, to make the new scaling-up method more understandable for users, the framework is detailed and interpreted in a more mathematical way, shown in Figure 2.4. Most steps remain unchanged, and only mathematical explanations are added in order to further explain the steps. The step of collecting data in the framework is divided into two steps: (a) collecting data in part of local scenarios as sample data (both of the data of indicator values and the data of situational variable values) and (b) collecting data in the rest of local scenarios (only the data of situational variable values).
2.1. DESIGN OF THE NEW SCALING-UP METHOD

Figure 2.3: The graphic interpretation of the new scaling-up method

1. A large scenario including a number of small scenarios
2. Know a few small scenarios in the large scenario
   Known situational variables and impacts in a few small scenarios → Relationship
3. From a few to all
   Scale sideways
   Known situational variables (in the rest of small scenarios) → Relationship (from 2) → Impacts in the rest of small scenarios
4. Aggregate the impacts in small scenarios to the large scenario
A: the set of all situational variables, \( \{sv_1, sv_2, sv_3, \ldots, sv_v\} \)

N: the set of numerical situational variables, \( N \subseteq A \) and \( N \neq \emptyset \), \( \{nsv_1, nsv_2, nsv_3, \ldots, nsv_v\} \)

C: the set of categorical situational variables, \( C \subseteq A \), \( \{csv_1, csv_2, csv_3, \ldots, csv_v\} \)

\[ Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_v \end{bmatrix}, \quad NSV = \begin{bmatrix} nsv_1 \\ nsv_2 \\ \vdots \\ nsv_v \end{bmatrix} \]

\( y_i \): the value of the indicator of the impact in small scenario \( i \)

\( nsv_i \): the value of numerical situational variable \( nsv_i \) in small scenario \( i \)

\( i = 1, 2, 3, \ldots, s ; j = 1, 2, 3, \ldots, p \)

Figure 2.4: The mathematical interpretation of the new scaling-up method
2.2. INSTRUCTION FOR USING THE NEW SCALING-UP METHOD

The framework of the new scaling-up method shows the main procedure of assessing the impacts in the large scenario by scaling up, while how to proceed each step is not shown in it. Therefore, in this section, some advices (examples) for how to proceed each step are proposed for better applying the new scaling-up method in practice.

2.2.1. DETERMINING A SUITABLE INDICATOR

In fact, when evaluating the impacts of an ITS application on traffic efficiency, defining an indicator of the impacts is to some extent defining a measure of traffic efficiency, because the indicator is basically the change of the measure of traffic efficiency after adopting the ITS application. The common-used way to choose a suitable measure of traffic efficiency is literature study. In the previous studies, there are many cases using various measures of traffic efficiency under different situations. These measures could be the alternate options. Only one measure would be chosen in one scaling-up case, so in order to make a sensible choice among all alternate measures, a set of criteria should be defined. The criteria could be based on the characteristics of the studied ITS application, the features of the scenarios, the opinion of policy makers, and so on. According to the criteria, the suitable measure of traffic efficiency can be selected among the measures mentioned in literature, and the indicator of the impacts of the studied ITS application can be defined with the change of this measure after adopting the ITS application.

The indicator could also be directly decided by the opinion of policy makers. For example, if the policy makers want to know the influence of one ITS application on the total travel time specifically, the indicator should be the change in total travel time after adopting the ITS application. Another valuable reference might be the previous cases using the new scaling-up method. For example, if in a study which assesses the impacts of one ITS application by the new scaling-up method, the change in average travel time is chosen as the indicator of the impacts, it can be directly adopted in other researches that also study the same ITS measure if needed.

2.2.2. CHOOSING SUITABLE SITUATIONAL VARIABLES

The process of choosing suitable situational variables could be similar to the process of determining the indicator of the impacts.

One possible way is choosing situational variables based on literature and the criteria. Since situational variables should have the largest influence on the impacts of the studied
ITS application [Jonkers et al., 2014], suitable situational variables should be influential to the effectiveness of the ITS application. Therefore, suitable situational variables can be selected based on the criteria among the influential factors (to the effectiveness of the ITS application) mentioned in literature. The criteria could be based on the characteristics of the chosen ITS application, the features of the scenarios, the opinion of policy makers, etc.

Besides, situational variables used in the previous researches which use the new scaling-up method to study the same ITS application could be a good reference. Sometimes, if needed, the chosen situational variables in other similar studies can be directly applied.

Note that the number of the chosen situational variables in a study should be controlled within reasonable amount in order to avoid massive data collection and calculation, while the preferred number of the considered situational variables is anyway decided by the users of the new scaling-up method.

2.2.3. Building a Deterministic Relationship

One recommended approach to build a deterministic relationship is that assuming a mathematical function to represent the relationship and then conducting a regression analysis to examine if the assumed relationship is reliable and determine the parameters in the hypothetical function. How well the data match the hypothetical function can be represented by the value of R-squared. How accurately the hypothetical function estimates the local impacts can be indicated by a residual analysis. If the hypothetical function can estimate the impacts through the numerical situational variables with acceptable accuracy, it means that a deterministic relationship can be found.

There are various ways to assume the deterministic relationship properly. Users of the new scaling-up method can assume the relationship based on literature. A possible relationship can also be found by data analysis. The previous researches which use the new scaling method to study the same ITS measure could be a good reference as well.

Users of the new scaling-up method can apply but not limit themselves to the above recommended methods. After all, a proper way to build a deterministic relationship should depend on the characteristics of the collected data.

2.2.4. Aggregating the Local Impacts to the Large Scale

When aggregating the local impacts to the large scenario, the way of aggregating mainly depends on the characteristics of the indicator of the impacts. For example, if the indicator is travel time per cap, the population in each small scenario might be taken into account to
make a sensible assessment of the impacts in the large scenario. But if the indicator of the impacts is total queue length, the impacts in the large scenario might just be the sum of the impacts in the small-scale scenarios.

Therefore, there is not a general aggregation method for all cases. Users of the new scaling-up method can decide an appropriate approach based on their specific situations.

2.3. CONCLUSION

In this chapter, the new scaling-up method is designed, which is expected to be able to assess the impacts of ITS with direct network-wide influences. But whether this new scaling-up method can work in practice is still unknown. So in the next chapter, this new scaling-up method is applied to a specific ITS for examining its applicability.

Besides, according to the design process of the new scaling-up method, setting requirements is quite helpful, which can indicate a clear view of what the new method should be like. Additionally, when designing a new scaling-up method, analysing the flaws and merits of the current approaches is also useful. The flaws could provide a perspective what is needed to be improved in the new method while the merits could be adopted directly in the new method. The design idea might be referential for the later designs of other methods.
3

APPLYING THE NEW SCALING-UP METHOD 
TO A SPECIFIC ITS APPLICATION

The new scaling-up method is applied to a specific ITS as a case study in this chapter. In Section 3.1.1, a specific ITS application (the on-trip dynamic navigation system) is chosen as the study object, additionally, the local scale and the large scale are defined. The indicator of the impacts of on-trip dynamic navigation and the situational variables are determined in Section 3.2 and 3.3. After that, Section 3.4 explains how the data is collected. Data analysis is conducted in Section 3.5 in order to build the deterministic relationship between the impacts and the situational variables. After that, Section 3.6 shows the specific scaling-up procedure for the on-trip dynamic navigation system. In the end, the conclusions are drawn for the whole chapter.

3.1. BEFORE APPLYING THE NEW SCALING-UP METHOD

Before applying the new scaling-up method to a specific ITS application, some decisions need to be made. On one hand, it is necessary to choose an ITS application as the study object. On the other hand, the local scale and the large scale need to be defined. If a project is supported by government (e.g. euroFOT), since the target large-scale scenario is decided by policy makers and the local impacts are assumed to be already known, so there is no need to specifically define the local scale and the large scale and the main work of the project is scaling up the local impacts to the target large-scale. However, the study conducted in this chapter is more like an academic research not a practical project, the local scale and the large scale are not naturally known thus need to be defined.
3.1.1. **Choice of the study object**

There are so many types of ITS working in the world. It is not reasonable to randomly choose one as the object of the case study. According to the research objective, the **basic requirement** which should be met by the chosen ITS application is that its **direct** influence should be **network-related** — the impacts are **directly** on the traffic assignment, the modal split, the departure time or the travel demand.

<table>
<thead>
<tr>
<th>ITS Application</th>
<th>Trip generation</th>
<th>Mode choice</th>
<th>Time choice</th>
<th>Route choice</th>
<th>Driving behaviour and vehicle conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traveller Information System</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Passenger Information</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>On-trip Dynamic Navigation</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Collective Re-routing System (e.g., DRIPs)</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Junction Control System (e.g., ramp metering)</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Electronic Toll Collection</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Adaptive signal control</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Road Section Control System (e.g., dynamic speed limits)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Autonomous Driving</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Cooperative Intersection Collision Avoidance System</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Lane Change Assistance System</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Intelligent Speed Adaptation/Assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Collision Warning System</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Adaptive Cruise Control (ACC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Cooperative Adaptive Cruise Control (CACC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Predictive Cruise Control (PCC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 3.1: Different ITS and their main /first-order impact level(s)

In Table 3.1, the existing popular ITS and their major /first-order impact level(s) are examined.

ITS applications which only have direct impacts on driving behaviour and vehicle conditions (e.g., Adaptive Cruise Control (ACC), Lane Change Assistance System and Collision Warning System) are excluded since they cannot meet the basic requirement in other words, their impacts are not directly on network level. In addition, it is not preferred to choose a ITS application whose impacts are on too many aspects (e.g. Traveller Information System) because analysing of the impacts might be too difficult. After excluding these two types of ITS, there are only five alternative ITS applications: **On-trip Dynamic navigation system**, **Collective Re-routing System**, **Junction Control System**, **Electronic Toll Collection and Adaptive Signal Control**. Furthermore, considering the impacts of **Collective Re-routing System**, **Junction Control System**, **Electronic Toll Collection and Adaptive Signal Control** are closely related to the installation sites of these ITS applications since their major impacts are on the traffic around their locations, the **on-trip dynamic navigation system** is chosen as the particular ITS for the case study.
3.1.2. LOCAL SCALE V.S. LARGE SCALE

As mentioned in Section 1.1.3, the large scale often represents country level or EU level. So the local scale should be on the lower level (e.g. state level or city level). Considering the on-trip dynamic navigation system can improve traffic efficiency mainly by changing the route choice, the local scale should at least be a network level and the corresponding networks (local scenarios) should have enough roads in order to provide multiple alternative routes for drivers. But the corresponding networks (local scenarios) should not be too large to make data collection difficult.

In this chapter, the local scale is defined as urban network level. On one hand, data collection conducted in a city is not quite difficult. If simulations are conducted to collect data, building a model for a city (except the mega city like London, Beijing or Paris) is still feasible. On the other hand, urban networks often have high density so that the drivers could have various alternative routes. The on-trip dynamic navigation system could have significant impacts on traffic efficiency in urban networks.

Note that if urban networks are chosen as the local scenarios, when scaling up the local impacts to the large scale, the impacts of the on-trip dynamic navigation system on inter-city roads cannot be taken into account. However, since urban networks are much denser than intercity roads thus can provide much more alternative routes, the on-trip dynamic navigation system can make much more significant influences on urban networks than on intercity roads, therefore, the ignorance of the impacts on intercity roads is acceptable.

Besides, the large scale could be country level or EU level. Considering the entire EU is quite large compared to a city, the country level is preferred. However, if needed, the large scale can still be EU level although more local impacts might need to be known.

3.2. DETERMINATION OF THE INDICATOR OF THE IMPACTS ON TRAFFIC EFFICIENCY

Before investigating the relationship between the impacts and the situational variables, the indicator of the impacts on traffic efficiency needs to be determined.

According to Section 2.2.1, it is recommended to choose the suitable measure for traffic efficiency before determining the suitable indicator of the impacts. Therefore, in this section, the measure of traffic efficiency is chosen firstly. The indicator of the impacts of the on-trip dynamic navigation system on traffic efficiency is defined as the change of the measure after adopting the on-trip dynamic navigation system. Additionally, there are neither referential researches (i.e. studies which use the new scaling method to assess the
Applying the New Scaling-up Method to a Specific ITS Application

Impacts of the on-trip dynamic navigation system) nor intervenes of political opinions, so the recommended approach for determining the measure of traffic efficiency in Section 2.2.1, which sets criteria and selects the suitable measure among the measures mentioned in literature, is used.

Note that in this chapter, the impacts of the on-trip dynamic navigation system represent the impacts on all drivers, including informed drivers (i.e. drivers with on-trip dynamic navigation) and uninformed drivers (i.e. drivers without on-trip dynamic navigation), because it is assumed that policy makers would care all involved drivers not only informed drivers. Additionally, the effectiveness of the on-trip dynamic navigation system indicates the positive impacts of the on-trip dynamic navigation system on traffic efficiency.

3.2.1. Literature Review: Measure of Traffic Efficiency

This section reviews the previous related studies to find out the alternative measures of traffic efficiency for the later selection of the suitable measure. In addition, some former researches also set some criteria when chose the measures of traffic efficiency. These criteria could be a reference for the setting of criteria in choosing the suitable measure of traffic efficiency.

In a former report [Kaparias et al., 2011] which tried to determine the key performance indicators for ITS, traffic efficiency was constituted by four sub-categories: mobility; reliability; operational efficiency; and system condition and performance. Additionally, specific measurements for each sub-category were listed in that report, like travel time, delay, commuting time, etc.

But mostly, traffic efficiency is represented by severity (magnitude) of congestion in the network.

One fact is most researches studying the impacts of ITS on traffic focused on congestion-related measures such as saved travel time [Sparmann, 1991], average travel time [Levinson, 2003; Toledo and Beinhaker, 2006], delay [Faber et al., 2012] and level of service (LOS) [Neth et al., 2005].

Furthermore, to better evaluate congestion, former studies also classified or/and assessed different congestion measures via different perspectives. For instance, Bertini [Bertini, 2005] stated that there are four general categories of congestion measures based on described objectives: i) measures that explain the duration of congestion experienced by users (delay, risk of delay, average speed, and travel time, etc.); ii) measures that analyse how well the system is functioning at a given location (volume to capacity (V/C) ra-
3.2. determination of the indicator of the impacts on traffic efficiency

3.2. Determination of the Indicator of the Impacts on Traffic Efficiency

Aftabuzzaman divided congestion measures into four broad groups due to their forms [Aftabuzzaman, 2007]: i) basic measures like total delay and number of congested travel; ii) ratio measures such as travel rate, delay rate and delay ratio; iii) LOS; and iv) indices for example congestion index and travel rate index. Furthermore, in a recent study, Rao created a different classification of congestion measures, using metrics of speed, travel time, delay, volume and level of service [Rao and Rao, 2012]. Beyond congestion measures mentioned above, there are also researches measuring congestion levels through drivers’ feeling [Thianniwet et al., 2009; Ye et al., 2013].

Not only on the categories of congestion measures, researches held diverse opinions but also on the question of “what is the best (most suitable) congestion measure?”.

In a American study which tried to define a performance measure(s) that could be used to indicate congestion levels on critical corridors, Medley and Demetsky reviewed measures used in the literature [Medley and Demetsky, 2003]. Based on the analysis of these measures, the study in the end chose total delay and buffer index as the performance measures and listed the reasons why they were chosen while others were not (see details in Figure A.1). According to the conclusion drawn in this research, both of total delay and buffer index are useful as an evaluation tool for congestion. Particularly, the study insisted that total delay is suitable for use by transportation professionals because it applies to all vehicles and drivers, while buffer index is appropriate for the public because it addresses individual travel time. But it should be noted that the main criterion to choose measures in that study was data availability, so the choice made can offer a good recommendation but cannot be generally adopted.

In another research, Bertini nonetheless made a general discussion of definitions of traffic congestion and ways it is measured by a survey among transportation professionals [Bertini, 2006]. In that survey, respondents mentioned time, speed, volume, LOS and traffic signal cycle failure as primary definitions of congestion when ‘speed’, ‘volume’ and ‘time’ were the three most often cited. Certainly it is not the whole story, in practice, the definition of congestion might depends on the measured area. For motorways(highways) speed is more likely to be measured meanwhile traffic signal cycle failure is considered often in urban areas. Additionally, these professionals involved assumed delay was the most popular congestion measure followed by LOS, speed, travel time and volume to capacity (more details could be found in Figure A.2).

Aftabuzzaman [Aftabuzzaman, 2007] otherwise provided a comprehensive critique of traffic congestion measures from another perspective. Based on literatures, Aftabuzzaman proposed six criteria for assessing congestion measures: i) demonstrates clarity and sim-
Applying the new scaling-up method to a specific ITS application

According to Bertini and Aftabuzzaman, delay appears as a good measure for congestion. However, Rao gave a total different estimate to delay [Rao and Rao, 2012]. In that study, Rao created a new set of criteria: i) simplicity; ii) ease of data collection; iii) stability; iv) repeatability; v) magnitude of congestion; vi) city comparison; and vii) continuous value. According to the conclusion, Rao and Rao believed that congestion measurement criteria based on speed should be adopted because traffic speed is highly sensitive parameter. One the other hand, delay-related measures cannot meet the most of the criteria. But unfortunately, the conclusion drawn in the study (Figure A.4) is quite rough. The authors did not provide any detailed explanations about how they decided whether one measure can meet one criterion thus the judgement is very doubtable actually. The most obvious evidence of the hush conclusion is that the authors lowly valued the delay measure but at the same time inferred that total delay could allow transportation professionals to estimate how improvements within a transportation system affect a particular corridor or the entire system.

Litman described various measures for congestion based on literature [Litman, 2013]. According to Litman, some of these measures only measure vehicle traffic delay at a particular location while others are more comprehensive. The complete list of measures and evaluation is shown in Figure A.5. Besides, Litman also drawn some recommendations for measuring congestion costs. He insisted that congestion intensity measures, such as roadway level-of-service and the travel time index, are useful for making short-term decisions, such as how best to travel across town during rush hour, but are unsuitable for strategic planning decisions that affect congestion exposure, the amount that travellers must drive under urban-peak conditions.
3.2.2. CRITERIA FOR CHOOSING THE MEASURE OF TRAFFIC EFFICIENCY

Referring to the previous researches and considering the characteristics of the on-trip dynamic navigation system, three criteria for choosing the suitable measure of traffic efficiency are set as follows. The reason for setting each criterion is also shown.

1. **Continuous value**
   
   Since the approach used in this study tries to find out the deterministic relationship between the situational variables and the indicator (i.e. the change of the measure after adopting the on-trip dynamic navigation system), so the measure should have continuous values.

2. **Ease of data collection on network level**

   Because the impacts of the on-trip dynamic navigation system on traffic efficiency are network-wide, the measure should be able to be calculated on network level. In addition, available time spent on the data collection is quite limited, so the chosen measure should be able to be collected easily.

3. **Absolute reflection of traffic efficiency**

   A good measure should be able to reflect the *absolute* traffic efficiency. In other words, the difference in the value of the measure can definitely indicate the difference in traffic efficiency. For example, when the value of the measure in network A is higher than that in network B, the traffic efficiency in network A is definitely higher than that in network B. On one hand, the ability of the measure to reflect the absolute traffic efficiency can make the comparison among the networks meaningful — which network has the highest traffic efficiency could be directly determined by the values of the measure. On the other hand, this ability can make the indicator of the impacts (i.e. the change of the measure after adopting the on-trip dynamic navigation system) to be able to precisely show the impacts of the on-trip dynamic navigation system on traffic efficiency — the increased measure means that traffic efficiency is improved while the reduced measure indicates that traffic efficiency is decreased by the on-trip dynamic navigation system.

3.2.3. CHOICE OF THE MEASURE OF TRAFFIC EFFICIENCY

The full check list is shown in Table 3.2, also including the alternative measures and their definitions. Note that in the table, **Y** means the measure can meet the criterion while **N** means the measure cannot meet the criterion. Only when all three requirements are met by the measure, the column of *Check result* would be filled with **Y**.
It is not difficult to judge whether a measure has continuous values. It is found that all mentioned measures except Level of Service have continuous values.

Whether a measure is easily collected on network level mainly depends on how the measure is calculated. The speed, the travel time, the queue length can be easily collected on network level and the related data are often available in many areas, therefore, measures which are based on these elements (e.g. average speed, average travel time, delay and queue length) can also be easily collected on network level. However, the data of certain parameters (e.g. commuting time, acceptable travel time and time cost in congested conditions) requires much efforts to be collected and is often not available, then measures which need to be calculated via these elements (e.g. Congestion Burden Index, Delay Rate and Travel Rate Index) cannot meet the second criterion.

To judge if a measure is able to reflect the absolute traffic efficiency, one way is investigating what factors can influence the measure. A measure which highly depends on its specific circumstances often cannot indicate the absolute traffic efficiency. For example, average travel time, which mostly is greatly affected by the local speed limit and the trip distance, is not suitable for this study. Likewise, Average Speed (influenced by the speed limit), Total Delay (influenced by the number of travellers), Queue Length (influenced by the total length of roads) and Congested travel (influenced by the number of travellers) are not desirable measures either. Furthermore, although a certain measure is not significantly affected by its specific circumstances, its still cannot reflect the absolute traffic efficiency, because the change in traffic efficiency cannot be indicated in the change of the measure. For example, it is totally possible that even if the traffic efficiency is improved, the Congested Time keeps unchanged (e.g. the duration of the morning peak is often the same), so Congested Time cannot reflect the absolute traffic efficiency.

According to Table 3.2, only two measures that can meet all requirements: Speed Reduced Index and Travel Time Index. Compared to Speed Reduced Index, Travel Time Index is more directly related to the travel time thus the travel costs. Considering the impacts of ITS are often used for a Cost-Benefits Analysis, Travel Time Index (TTI), average travel time divided by average free flow travel time, is chosen as the measure of traffic efficiency.

3.2.4. Definition of the Indicator of the Impacts of the On-Trip Dynamic Navigation System

The indicator of the impacts is naturally defined with the change of the measure of traffic efficiency after adopting on-trip dynamic navigation system. But the change of TTI sometime cannot indicate the effectiveness of the on-trip dynamic navigation system precisely.
### 3.2. Determination of the Indicator of the Impacts on Traffic Efficiency

#### Table 3.2: Assessment of the measures of traffic efficiency based on the criteria

<table>
<thead>
<tr>
<th>Category</th>
<th>Indicator</th>
<th>Definition</th>
<th>Continuous value</th>
<th>Ease of data collection on network level</th>
<th>Absolute reflection of traffic efficiency</th>
<th>Check result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>Average Speed</td>
<td>Average vehicle travel speeds</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Speed Reduction Index</td>
<td>Ratio of the decline in speeds from free flow conditions</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Travel Time</td>
<td>Average Travel Time</td>
<td>Average trip time</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Buffer Index</td>
<td>Extra time (buffer) needed to ensure on-time arrival for most trips</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Congestion Burden Index</td>
<td>Travel rate index multiplied by the proportion of commuters subject to congestion by driving to work</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Planning time index</td>
<td>Extra time most travellers include when planning peak period trips</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Travel Rate Index</td>
<td>Ratio of time to travel in congested conditions than in uncongested conditions</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Travel Time Index</td>
<td>Average travel time divided by average free flow travel time</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Delay</td>
<td>Congestion Index</td>
<td>Ratio of link delay (the difference between actual and acceptable travel time) to acceptable travel time</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Congestion Severity Index</td>
<td>Delay per million miles of travel</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Delay Rate</td>
<td>Actual travel rate (minutes per mile) minus acceptable travel rate (minutes per mile)</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Delay ratio</td>
<td>Delay rate divided by acceptable travel rate (minutes per mile)</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Total Delay</td>
<td>Sum of time lost for all vehicles</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Queue</td>
<td>Queue Length</td>
<td>The length of queue</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>LOS</td>
<td>Level of Service</td>
<td>Intensity of congestion at a particular roadway or intersection, rated from A (uncongested) to F (most congested)</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Other</td>
<td>Congested Time</td>
<td>Hours of congestion exist</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Congested travel</td>
<td>Amount of travel that occurs during congestion</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Percent of Congested Travel</td>
<td>The congested vehicle-miles of travel divided by total vehicle-miles of travel</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

The table above presents various measures of traffic efficiency along with their definitions, continuous value (Y for continuous, N for non-continuous), ease of data collection on network level, and absolute reflection of traffic efficiency. The check results are also indicated.
3. Applying the New Scaling-up Method to a Specific ITS Application

Table 3.3 shows an example. Note that the column *Before* shows the average travel time before adopting the on-trip dynamic navigation system and the column *After* represents the average travel after adopting the on-trip dynamic navigation system. For case 1 and case 2, the change of TTI is the same, but on common sense, the effectiveness of the on-trip dynamic navigation system in case 2 seems higher since the original travel time is reduced relatively more.

<table>
<thead>
<tr>
<th>Case</th>
<th>Average free flow travel time (min)</th>
<th>Before (min)</th>
<th>After (min)</th>
<th>Change of TTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>50 (TTI=5)</td>
<td>45 (TTI=4.5)</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>30 (TTI=3)</td>
<td>25 (TTI=2.5)</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 3.3: An example of the inability of the change of TTI to indicate the impacts of the on-trip dynamic navigation system

Therefore, the indicator of the impacts of the on-trip dynamic navigation system is defined with the relative change of TTI, as follows\(^1\) \(^2\).

\[
\text{Relative Reduced TTI} = \frac{\text{TTI (before)} - \text{TTI (after)}}{\text{TTI (before)}} \times 100\% \quad (3.1)
\]

According to the definition, the higher the Relative Reduced TTI is, the more effective the on-trip dynamic navigation system is.

3.3. Choice of the Situational Variables

Similarly to the choice of the indicator of the impacts of the on-trip dynamic navigation system, since there are no referential researches, the recommended approach in Section 2.2.2, which sets criteria and chooses the situational variables among the influential factors of the effectiveness of the on-trip dynamic navigation system, is used in this section.

It is firstly needed to be known how the effectiveness of the on-trip dynamic navigation system could be influenced, in other words, what factors can affect the effectiveness of the on-trip dynamic navigation system. The on-trip dynamic navigation system can influence the effectiveness of the on-trip dynamic navigation system mainly through changing the route choice (i.e. traffic assignment) and thus making the whole network to be more or less equally loaded [Kaparias et al., 2011]. Therefore, it could be deduced that the factors influencing the route choice should be able to affect the effectiveness of the on-trip dynamic navigation system.

\(^1\)TTI (before): TTI before adopting the on-trip dynamic navigation system.
\(^2\)TTI (after): TTI after adopting the on-trip dynamic navigation system.
3.3. CHOICE OF THE SITUATIONAL VARIABLES

3.3.1. LITERATURE REVIEW: INFLUENTIAL FACTORS OF THE EFFECTIVENESS OF THE ON-TRIP DYNAMIC NAVIGATION SYSTEM

Chen et al. developed an individual-based mechanism to find the crucial criteria influencing drivers’ route choices [Chen et al., 2001]. Based on an empirical study in Taipei City, Chen et al. concluded travel time is the most important influential factor for route choice, followed by traffic conditions, travel time reliability, traffic safety, travel expenses, drivers’ familiarity, travel comfort, travel distance and driver habits in that order.

Similarly, in another research, Chang and Chen insisted that travel time is the most significant factor while number of intersections, traffic safety, traffic lights, habits, cognitive skills and other behavioural considerations also play roles [Chang and Chen, 2005].

In addition, Pan stated that travel time difference is very important for switching route under a bi-route situation [Pan and Khattak, 2008] in one experiment.

Furthermore, another experiment shown that drivers prefer a route that has lower travel time, higher speed and fewer number of stops and also more likely to choose routes that are efficient, easy to drive and familiar given real-time information [Zhang and Levinson, 2008]. Additionally, trip purpose also plays an important role in that experiment: when making commute, event, and visit trips, drivers tend to choose a familiar and reliable route under the time pressure.

In the master thesis of Kemunto [Oyaro, 2013], a conclusion of influential factors on route choice is also drawn. Travel time and its reliability, traffic conditions, driver habits and cognitive skills, road characteristics, driver experiences and familiarity with the route were considered.

Bogers et al. nonetheless paid specific attention to travellers’ habit and learning [Bogers et al., 2005]. Based on laboratory experiments, Bogers et al. drawn a conclusion that learning about route attributes is especially important during the first days but then plays a smaller role than the provided information and the developed habit, and the way information is presented has a great impact on route choice.

Besides, Mahmassani et al. [Mahmassani et al., 2003] and Li et al. [Li et al., 2005] did studies about the impact of socio economic factors like age, gender and income on route choice.

Note that the mentioned factors above are driver-based, which can influence the individual driver’s route choice and thus affect traffic distribution in the network. But traffic assignment also depends on other network-level elements, like the number of alternate routes and the network traffic condition.
High amount of alternate routes can offer drivers abundant choices and thus possibly stimulate drivers to switch to ‘better’ routes. The number of alternate routes is mainly based on the layout or design of the road network [Muizelaar, 2011]. The design of the network often determines the amount of junctions and traffic lights, one-way roads, speed limits, and so on. Frejinger, for instance, stated that the amount of left-turns played a significant role in his route choice model [FREjINGER, 2008].

On the other hand, network traffic condition also plays an important role. If network is highly congested, it is difficult for travellers to have a better route with a shorter travel time.

Furthermore, information characteristics are mentioned in literature as well. Many studies set the percentage of informed drivers (penetration of ITS) as an essential variable when studying the impacts of traffic information or ITS [Mahmassani and Jayakrishnan, 1991; Emmerink et al., 1995; Levinson, 2003; Pan and Khattak, 2008; Gao, 2012; Balakrishna et al., 2013].

In one study, Balakrishna cited the influence of information update interval on effectiveness of ITS [Balakrishna et al., 2005]. In addition, Muizelaar stated information resource could also affect the route choice. For example, a Variable Message Sign (VMS) provides different information with an individual in-car device. VMS can offer the information in some locations while navigation systems can provide continuously traffic information. Muizelaar also believed that the quality, cost and types of information are significant as well and cited some reference like [Srinivasan and Mahmassani, 2003; Jou et al., 2005; Chorus et al., 2006].

### 3.3.2. Criteria for Choosing Situational Variables

Based on the characteristics of the on-trip dynamic navigation system, three criteria for choosing situational variables are made and why the criteria are set is explain as follows.

1. **Numerical**

   Although categorical situational variables could be used, more efforts are required to adopt categorical situational variables since different deterministic relationships have to be built in various categories. Therefore, in this study, only numerical situational variables are chosen at the beginning. If categorical situational variables are found necessary in the later analysis, they would be considered then.

2. **Ease of data collection on network level**

   The data of situational variables could be easily found or collected on network level in order to save the data collection time.
3. **Influential on network level**

The local scale is defined as urban network level, so the impacts of the on-trip dynamic navigation system should be considered on network level (i.e. on all travellers) instead of on individuals. The chosen situational variables should be able to influence the effectiveness of the on-trip dynamic navigation system on network level, and individual-based influential factors are not necessarily to be considered.

3.3.3. **Choice of the situational variables**

According to literature, the mentioned influential factors of the effectiveness of on-trip dynamic are divided into five categories. The full check list is shown in Table 3.4. Note that in the table, Y means the measure can meet the criterion while N means the measure cannot meet the criterion. Only when all three requirements are met by the measure, the column of *Check result* would be filled with Y.

It is not difficult to judge if an influential factors could be a numerical situational variable. For example, the driver’s habit and driver’s experience are difficult to be represented in values thus could not be numerical situational variables.

In addition, influential factors which cannot be represented in values often are difficult to be collected on network level, because they contains too detailed information (e.g. traffic light setting) or too personal (e.g. trip purpose). Except the individual-based influential factors, most mentioned factors are considered influential on network level.

According to Table 3.4, *network density, intersection density, congestion level, penetration of the on-trip dynamic navigation system* and *information update interval* are found to be able to meet all requirements.

However, considering the intersection density is often consistent with the network density, in other words, the higher network density often indicates the higher intersection density. Therefore, **to avoid the dependence between situational variables, network density is chosen as a situational variable while intersection density is not**, since the collection of network density is easier compared to that of intersection density. In conclusion, the following (numerical) situational variables are selected for this study:

- Network density: the average length of roads in 1 square kilometre
- Congestion level
- Penetration rate: the penetration of the on-trip dynamic navigation system
- Information update interval: update interval of on-trip information provided to drivers
<table>
<thead>
<tr>
<th>Category</th>
<th>Influential factors</th>
<th>Numerical</th>
<th>Ease of data collection on network level</th>
<th>Influential on network level</th>
<th>Check result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trip purpose</td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>familiarity to network</td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>driver’ experiences</td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>drivers’ habit</td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>drivers’ cognitive skills</td>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>socio-economic factors</td>
<td>(gender, age, income, etc.)</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Network geometry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>network density</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>number of intersection</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>traffic lights setting</td>
<td></td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Routing characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>penetration rate</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>information update interval</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>information resource</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>quality of information</td>
<td></td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Demand &amp; Supply characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>congestion level</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 3.4: Assessment of situational variables based on the criteria
3.3. Choice of the situational variables

3.3.4. Measures of the situational variables

The measures of the network density, the penetration rate and the information update interval are quite clear based on their definition, so it would not be discussed further in this section. The measure of the congestion level needs to be determined. As mentioned in Section 3.2, Travel Time Index (TTI) is a good indicator for congestion, so when measuring the congestion level, TTI could also be used. Nonetheless, to make the measure more sensible, TTI is improved to a new measure — **Congestion Level Factor (CLF)** (the deification is shown in Equation 3.2) is created. When the value of CLF is zero, it means there is no congestion. The higher the Congestion Level Factor is, the more severe the congestion is.

\[
CLF = TTI - 1
\]  

(3.2)

Note that Travel Time Index (TTI) is indirectly used as situational variable (in the congestion level factor) as well as used in the measure of the effectiveness (to compute the Relative Reduced TTI). But this would not lead to a necessary correlation between the congestion level and the effectiveness of the on-trip dynamic navigation. The current TTI is used to represent the congestion level while the relative change in TTI due to the on-trip dynamic navigation system is used to show the effectiveness.

3.3.5. Hypotheses of the influences of the situational variables

Based on literature, the hypotheses of the influences of the four situational variables can be made, which might be referential for build the deterministic relationship:

- **Hypothesis 1: Network density**
  
  The network density should have a positive relationship with the effectiveness of the on-trip dynamic navigation system. Higher network density means more roads per square kilometre thus more alternative routes. Higher amount of alternative routes can provide more route choices for drivers, so the traffic efficiency can more likely be improved by guiding drivers to ‘better’ routes. Therefore, it is expected that higher network density can result in higher effectiveness of the on-trip dynamic navigation system.

- **Hypothesis 2: Congestion level**
  
  When there is no congestion or slight congestion, the effectiveness of the on-trip dynamic navigation system would increase with the increasing congestion level. Because when the flow is much lower than the capacity and in the absence of incidents,
on-trip dynamic navigation has few opportunities to save time. When it is heavy congested, the effectiveness of the on-trip dynamic navigation system would decrease with the increasing congestion level since uncongested alternative routes may not be available either [Wunderlich, 1995].

- **Hypothesis 3: Penetration rate**

  When the penetration rate is low, the penetration rate has positive relationship with the effectiveness of the on-trip dynamic navigation system because the on-trip dynamic navigation system can lead informed drivers to ‘better’ routes. But when the penetration rate reaches a saturation value, it has negative relationship with the effectiveness of the on-trip dynamic navigation system due to overreaction [Balakrishna et al., 2013]. (Overreaction often occurs when too many drivers respond to the information. Under this situation, a congestion may be transferred from the original area to the alternative routes since too many people follow the route guidance and change their routes [Bottom et al., 1999; Wahle et al., 2000; Ben-Akiva et al., 1991].)

- **Hypothesis 4: Information update interval**

  The shorter information update interval is, the more timely drivers can switch to ‘better’ routes [Balakrishna et al., 2013]. So it could be deduced that the on-trip dynamic navigation system with shorter information update interval could lead higher effectiveness.

### 3.4. **DATA COLLECTION**

This section mainly explains what data is collected and how it is collected.

Section 3.4.1 states that why simulations are used to collect data and why DYNASMART-P is chosen as the simulation tool. The series of data that needed to be collected are introduced in Section 3.4.2 and thus the range of values of the situational variables is declared in Section 3.4.3. Additionally, Section 3.4.4 explains some primary settings made in DYNASMART-P before and during the simulations. After that, some pilot simulations are conducted in Section 3.4.5 to primarily investigate the influences of the situational variables. At last, the collected data series are listed in Section 3.4.6.

#### 3.4.1. **DATA RESOURCE**

This section provides the reasons that why simulations are used to collect data and why DYNASMART-P is chosen as the simulation tool. Due to the lack of suitable practical data, simulations are chosen as the tool to collect data. Besides, mesoscopic simulation models
are considered as the best fit for this thesis compared to microscopic simulation models and macroscopic simulation models, based on the trade-off between efficiency and sufficiency of details. Among the current mesoscopic simulation models, DYNASMART-P is chosen as the simulation tool in this study considering the availability of software certification.

3.4.1.1. THE REASON OF USING SIMULATION DATA

Lack of suitable practical data  
Field tests are often considered to be good data resources when studying the impacts of ITS in reality. However, it is quite difficult to get enough practical data for on-trip dynamic navigation. Firstly, reliable data sets should be based on the same or similar dynamic navigation system. However, in real practice, various on-trip dynamic navigation systems are being used with different real-time data, different algorithms and different strategies of guidance. So although there exist some field tests of dynamic navigation, they can not be integrated and used in one study. Besides, in practice, traffic efficiency is influenced by lots of factors. Not every field test can guarantee that the result is not influenced by other external elements except for the dynamic navigation so the real-life data is hard to reflect the impacts of one single system.

Better control of collected data  
In the simulation, basically, expect for the limitation of simulators, researchers could obtain any data they want with reasonable scale. For example, the effectiveness of on-trip dynamic navigation under different demands in the same network could be collected in simulations but in reality only one demand (maybe with small daily variations) would exist. In other word, data collected from simulations is much richer than that collected in practice, which is quite valuable for research.

Therefore, data used in this chapter is collected by simulations.

3.4.1.2. TYPE OF SOFTWARE

As to choose a suitable simulation software, the first question is "which type of software could simulate on-trip dynamic navigation, macroscopic, microscopic or mesoscopic software?".

Macroscopic simulation tools are firstly considered because they can simulate a large network quickly. But at last Macroscopic tools are excluded because their inability / difficulty to adopt on-trip dynamic navigation strategy. In macroscopic simulation software, (generally) drivers are assigned routes before starting their trips and would not change their routes during the trip, although iterative simulations could be used to make the assignment
better. But the main function of on-trip dynamic navigation is dynamically planning routes during the trip to find a real-time optimal route. So macroscopic simulation software basically cannot simulate the on-trip dynamic navigation and thus isn't suitable for this study.

Microscopic software like ITS modeller could indeed simulate on-trip dynamic navigation properly. In microscopic simulation tools, drivers are considered separately — every driver has his/her own attributes, so the access to real-time data, the reaction to real-time data and the switch behaviour of each driver could be simulated accurately. However, to simulate an urban-scale network, microscopic simulation models require too much computation time to be applicable. For instance, to evaluate the impacts of on-trip dynamic navigation for a simple network, at least a one-day simulation is expected not including the model building time and debugging time. Therefore, considering that there is not much time for data collection in this study, microscopic simulation software is excluded as well.

Mesoscopic software combines the characteristics of both microscopic and macroscopic software. The traffic flow is represented by individual vehicles (like microscopic models), so the interaction between drivers and the impacts of information on individual drivers could be simulated. On the other hand, the movement of vehicles is governed by the average speed on the link (like macroscopic models) thus the computation time will not be too long. In summary, mesoscopic software can simulate the impacts of the on-trip dynamic navigation system within acceptable time consumption.

Therefore, mesoscopic simulation models are the best fit for this thesis based on the trade-off between efficiency and sufficiency of details.

3.4.1.3. DYNASMART-P

Mesoscopic models, compared to microscopic and macroscopic models, are not quite popular in the field of traffic simulation. They are mainly developed and applied in the United States. To the author’s knowledge, there are several well-known mesoscopic models in the world, which are Aimsun, AnyLogic, Cube Avenue, DYNASMART and DYNAMIT. But after checking the characteristics of these mesoscopic models, only DYNASMART and DYNAMIT could model the on-trip dynamic navigation system easily. On the other hand, DYNASMART and DYNAMIT are quite similar and developed by the same group. Finally, considering the availability of software certification, DYNASMART-P is chosen as the simulation tool in this study.

DYNASMART-P (an acronym for ‘dynamic network assignment simulation model for advanced road telematics’) is a dynamic traffic simulation model that was first developed by Peeta and Mahmassani in 1992. It is used to address complex and dynamic transporta-
tion operations and planning issues, particularly in the ITS context. DYNASMART-P overcomes the limitations of traditional static assignment and simulation models by using advanced traffic modelling techniques to capture the dynamics of congestion formation and dissipation associated with time varying demands and network conditions. It represents a new generation of traffic analysis tools to support transportation network operations and planning decisions. The modelling features chosen for implementation of DYNASMART-P achieve a balance between representation detail, computational efficiency, and input data requirements.

More information about DYNASMART-P could be found in Appendix B.

**Route choice in DYNASMART-P**

- If iterative simulation is used, drivers could be assigned routes based on user-equilibrium or system-equilibrium. But under such a case, drivers could not be divided in different classes — all drivers will follow user-equilibrium or system-equilibrium.

- If iterative simulation is not applied and just one simulation would be conducted, drivers could have four possible route choice modes.
  1. Drivers would be assigned routes based on external *path.dat* file which defines one and only one route for each selected vehicle. (Note that only iterative simulation could output *path.dat* file.)
  2. Drivers would randomly choose one available route at the departure and stick to this route during the trip.
  3. Drivers would have pre-trip information to choose the shortest route at the departure and stick to this route during the trip.
  4. Drivers would have en-route information and are able to switch their routes at any intersections if they want.

  Note: there are also several route choice modes related to VMS. But since VMS is considered in this chapter, they are not listed above.

**3.4.2. Date needed to be collected**

This section briefly explain what types of data are needed to be collected. The measures of the indicator of the impacts and the situational variables are discussed firstly and how the data of each measure could be collected is also stated.
3.4.2.1. Measures of related variables

Due to Section 3.2 and Section 3.3, there are five series of data needed to be collected: **Relative Reduced TTI** (Equation 3.1), **Network density**, **Congestion Level Factor (CLF)** (Equation 3.2), **Penetration rate**, and **Information update interval**.

The four situational variables should be able to be controlled by the simulation tool. The penetration rate and the information update interval can be directly determined by inserting needed numbers in the interface of DYNASMART-P. The CLF could be adjusted by changing the travel demand which is controlled by the *multiply* factor of the original demand matrix. The network density should be decided by the simulated network itself.

Additionally, the CLF and the Relative Reduced TTI need to be calculate by equations. According to Equation 3.2, to calculate the CLF, it is needed to know the current TTI. Besides, to calculate the Relative Reduced TTI, both of the TTI before adopting the on-trip dynamic navigation system and the TTI after adopting the on-trip dynamic navigation system are necessary to be known.

Furthermore, according to Section 3.2, TTI is defined with *average travel time divided by average free flow travel time*. Therefore, the actual collected data series from the simulations are *average travel time* and *average free flow travel time*.

3.4.2.2. Calculation of average free-flow travel time

To calculate the Relative Reduced TTI and the CLF, it is necessary to know the average free-flow travel time. However, the average free-flow travel time cannot be directly provided by the outputs of simulations. This section thus explain how the average free-flow travel time is calculated in this study.

Before the analysis, the mathematical terms used in the following content are explained as follows.

\[
\begin{align*}
v & \quad \text{vehicle } v, v \in V \text{ where } V \text{ is the set of all vehicles considered} \\
NV & \quad \text{number of all considered vehicles} \\
l & \quad \text{link } l, l \in L \text{ where } L \text{ is the set of all links in the network} \\
R_v & \quad \text{route chosen by vehicle } v, \text{ which is a set of all links passed by vehicle } v \\
LL_l & \quad \text{length of link } l \\
FFS_l & \quad \text{free-flow speed on link } l \\
FFT_l & \quad \text{free-flow travel time on link } l \\
FFT_v & \quad \text{free-flow travel time of vehicle } v \\
F_l & \quad \text{accumulative flow on link } l \text{ during the whole time period, in other words, the total number of vehicles passing link } l \\
AFFT & \quad \text{average free-flow travel time of all vehicles in } V
\end{align*}
\]
Based on the traffic flow theory, the free-flow travel time on link \( l \) could be expressed as Equation 3.3.

\[
FFT_l = \frac{LL_l}{FFS_l}
\]  

(3.3)

So the free-flow travel time of vehicle \( v \) could be expressed as Equation 3.4, and the average free-flow travel time of all considered vehicles could be expressed as Equation 3.5.

\[
FFT_v = \sum_{l \in R_v} FFT_l
\]  

(3.4)

\[
AFFT = \frac{\sum_{v \in V} \sum_{l \in R_v} FFT_l}{NV}
\]  

(3.5)

If Equation 3.5 is used to calculate the average free-flow travel time, then the route for every single vehicle needs to be known, which is a huge data set. So to simplify the calculation, Equation 3.6 instead of Equation 3.5 is chosen to compute the average free-flow travel time.

\[
AFFT = \frac{\sum_{l \in L} FFT_l \cdot F_l}{NV}
\]  

(3.6)

However, such an average free-flow travel time calculated by Equation 3.6, to some extent, can never be reached in a urban network in reality. Because, different from motorway networks, urban networks often include various intersections and waiting time at the intersections is unavoidable. When drivers pass a intersection, if the intersection is signalized, drivers can pass the intersection only when the light is green otherwise they have to wait for the green light; if the intersection is not signalized, drivers also have to wait sometimes to let the conflict traffic to pass. As a whole, the real average travel time is always higher than the average free-flow travel time calculated by Equation 3.6 although there is no so-called congestion.

Therefore, in this chapter, when the difference between the real average travel time and the average free-flow travel time calculated by Equation 3.6 is lower than 10% of the average free-flow travel time calculated by Equation 3.6, it is assumed the network has no congestion and the real average travel time is set as the \textit{Average free-flow travel time} and applied in Equation 3.1.
3.4.3. **Range of values of the situational variables**

In theory, there is no a compulsory range for values of the situational variables. But in practice, it is not feasible to cover all possible situations in one study. Therefore, the range of values of every situational variable is set as follows.

**Network density**  Although it is stated the networks should have enough roads in order to provide multiple alternative routes for drivers in Section 3.1.2, there is not a rigid standard for the network density and thus no range is set for values of the network density.

**Congestion level**  When the CLF is 1 (the actual average travel time is 2 times of average free-flow travel time), it could be assumed that the network is already quite congested. So in the following simulations, the range of the CLF is [0,1).

**Penetration rate**  Although the penetration rate is hard to reach 100% in practice, the scale of the penetration rate is set to from 0% to 100% in this study to better understand the influence of the penetration rate on the effectiveness of the on-trip dynamic navigation system.

**Information update interval**  Although a few systems use floating cars data to calculate the current traffic situation and predict travel time, most systems currently use detect loops to collect traffic data. It is known that generally loop data is sent to the control centre every 1 minutes. Adding data processing time, it can be assumed that real-time traffic information and route guidance could be updated at least every 2 minutes. On the other hand, it is thought 5-minutes interval is still acceptable for users of dynamic navigation. If a navigation can update information every 10 minutes, it is really too long. Thus the interested scale is from 2 minutes to 5 minutes.

3.4.4. **Primary settings in DYNASMART-P**

3.4.4.1. **Indifference band of switching behaviour**

According to last section, the indifference band and the threshold bound should be set for the switching behaviour. According to a research by Mahmassan [Mahmassani and Liu, 1999], the route switching indifference band for en-route path decision is about 18% on average and the threshold bound is around 1 minute required for a route switch. Thus in this study, the indifference is set to 18% while threshold bound is 1 minute.
3.4.4.2. Fraction of Informed Drivers that Responds to the On-Trip Dynamic Navigation System

In reality, although all informed drivers could have the en-route information, not everyone will follow the guidance. Part of informed drivers will continue their current routes although there might exist a ‘shorter’ route. However, since the response of informed drivers to information is not studied in this chapter, it is assumed that all informed drivers will switch to the shortest routes provided by the on-trip dynamic navigation system once both of indifference band and threshold bound are reached. Therefore, the fraction of informed drivers that responds to on-trip information is set to 1 in all simulations.

3.4.4.3. Use of Warm-up Period

To capture the effectiveness of the on-trip dynamic navigation system more precisely, warm-up time period is used in data collection. At the very beginning of simulation period, the vehicles in the network are very few which is not consistent with reality and thus travel time is always short. So if collecting the data from the beginning of simulation, travel time would be underestimated, which could lead to an inaccurate TTI. On the other hand, by running the simulation, it is found that average travel time quickly increases during the first 30 minutes and then keeps basically stable then. So the warm-up period is set to 30 minutes.

3.4.4.4. Route Choice Mode for Uninformed Drivers

Based on available route choice modes, uninformed drivers in DYNASMART-P needs to be assigned a route at the origin because they cannot switch routes during the trip. This original route could be a random route, from path.dat (as the output of another simulation) or based on pre-trip information.

Firstly, choosing a random route is too far away from reality, so it is not applied for uninformed drivers.

Then if the route choice is from another simulation, it would depend on how that simulation is run. As mentioned in many literature, the drivers’ familiarity to the network plays an important role in route choice behaviour. So the reference simulation could consider drivers’ familiarity to the network. Ideally, drivers are quite familiar with the network so basically they already choose a short route without the on-trip dynamic navigation system. But it seems very difficult to simulate ‘familiarity’ in DYNASMART-P. ‘Familiarity’ means drivers already know the approximate travel time on each route but certainly the real experienced travel time might be a little different from the estimated one. One possible way
to simulate ‘familiarity’ is using equilibrium assignment. But if drivers are assigned routes based on user or system equilibrium, they could get too good information beyond ‘familiarity’. Drivers could know the exact experienced travel time even before the trip. Certainly, we could adjust convergence threshold to output a not totally balanced equilibrium assignment. But under different demands in different networks, it is quite difficult to decide proper convergence threshold to ensure drivers in different situations have the same level of familiarity to the networks in order to make the results comparable. Thus, using equilibrium to represent the familiarity is not a good option. Another way might be using ‘historic data’ — the chosen route in the history. But unfortunately, such data is quite hard to find and even does not exist. Apart from the difficulty in simulation, in fact, not all drivers know the network very well in reality. One reason is that networks in many cities are quite complex and generally drivers can only be familiar with some roads but not all. Another reason is that with urban development, more and more drivers in the network don’t live in the city and are not familiar with the network. Therefore, getting route choice from another simulation seems not a smart choice.

On the other hand, pre-trip information is becoming more and more popular around the world. Local drivers would like to check the pre-trip information to have better route choice before departure although they are familiar with the network and exotic drivers even rely on pre-trip information to travel in an unfamiliar city. So it seems making drivers using pre-trip information to represent ‘uninformed situation’ is a good and practical option. Therefore, in the following simulations, the following assumption is applied:

*Uninformed drivers can have pre-trip information to choose the shortest route at the departure [All-or-nothing assignment] and would stick to this route during the trip; while informed drivers can change their routes during the trip due to real-time information.*

### 3.4.5. Pilot simulations

There are two main purposes of conducting pilot simulations.

One is to exclude the influence of external factors, which is achieved in Section 3.4.5.1. On one hand, lots of variables are able to influence the effectiveness of the on-trip dynamic navigation system but only few of them are chosen in this study. So only getting rid of the influence of unconcerned factors could make it possible to investigate the relationship between situational variables and impacts of ITS reliably. On the other hand, there possibly exist some elements in the simulation tool which can influence the result. These factors should also be eliminated.
Another aim is to primarily check if the chosen situational variables are really influential. Considering the amount of data that is collected in this section is quite massive, pilot simulations could avoid time waste on un-influential situational variables — if one situational variable is proved to be not crucial, adjustments could be made or this situational variable could be excluded in later simulations. The corresponding content is shown in Section 3.4.5.2, Section 3.4.5.3, Section 3.4.5.4 and Section 3.4.5.5.

3.4.5.1. Pilot 1: Influence of external factors

Tested network Network 1, one of test networks in DYNASMART-P, is chosen as the study object in pilot 1 (Map C.1). It is built based on the road network in Dallas Fort Worth, the United States. Network 1 is a typical American road network with a grid structure. There are about 181-kilometre roads over approximate 130 square kilometres in Network 1. Most of these roads are arterials with speed limit of 40 mph.

Influence of highways After a primary simulation of Network 1 in DYNASMART-P, it is found that most traffic flow is on the two highways across the network. This means that it is totally possible that the existence of highways can greatly influence the effectiveness of the on-trip dynamic navigation system, because even though the highways are quite congested they are still the best choice for most travellers. If one route is the best choice in most cases, the on-trip dynamic navigation system could be ineffective because few drivers have willing to switch their routes. To examine such a possibility, the impacts of highways are investigated and the result is shown in Figure 3.1. Note that two data series in Figure 3.1 are based on the same demand and the same information update interval (2 minutes). It could be seen that in Original Network 1, all values of Relative Reduced TTI under various penetration rates are close to zero indicating that the on-trip dynamic navigation system makes no influence in this situation. But when the two highways are deleted (data series Network 1 without highways), the values of the Relative Reduced TTI under various penetration rates become obviously positive, which reflects the impacts of the on-trip dynamic navigation system on traffic efficiency. Therefore highways could indeed influence the effectiveness of the on-trip dynamic navigation system significantly. But because the influence of highways is not considered as a study object in this chapter, so it is considered as an external factor. Since the influence of external factors should be excluded, in the networks that are used to collect data, there should be no highways.

Influence of actuated signals Generally, when studying the effectiveness of the on-trip dynamic navigation system, the impacts of other ITS should be excluded. But it is noted that most of traffic lights are actuated in Network 1 and actuated signals could be consid-
3. Applying the new scaling-up method to a specific ITS application

Figure 3.1: Influence of highways in Network 1 (Pilot 1-1)

Simulations therefore are conducted to investigate the influence of actuated signals and the result is shown in Figure 3.2. In Figure 3.2, *Network 1 without highways and actuated signals* means removing highways from Network 1 and transforming all intersections with actuated signals to unsignalized intersections, while *Network 1 without highways* means removing highways but keeping the actuated signals. Note that two data series Figure 3.2 are based on the same demand and the same information update interval (2 minutes). It could be found that there is an obvious difference between the two data series. In other words, actuated signals could have an influence on the effectiveness of the on-trip dynamic navigation system. Therefore, as mentioned previously, in the studied networks there should be no actuated signals.

**New tested network** Due to the strong influence of highways and actuated signals on the effectiveness of the on-trip dynamic navigation system, Network 2 (map C.2), which has similar network structure with Network 1 but does not include highways and actuated signals, is tested instead of Network 1 in the following content.

**Influence of random seeds** Random seeds exist in most simulation tools to initialize a pseudo random number generator. In DYNASMART-P, random seeds are mainly used to determine the vehicle loading procedure. Different random seeds could lead to different
3.4. DATA COLLECTION

Figure 3.2: Influence of actuated signals in Network 1 (Pilot 1-2)

Loading procedures thus different simulation results although the parameters are set as the same. So to make the data comparable, it is better to keep the same random seed in all simulations when collecting data from DYNASMART-P. However, if random seeds have significant influence on the result, using a fixed random seed would make the result unreliable. Therefore how significantly random seeds can influence the result need to be known in the first place. Simulations are thus conducted to check the influence of random seeds and the result is shown in Table 3.5. All data series in Table 3.5 are based on the same congestion level (CLF=0.19) and the same information update interval (2 minutes). According to Table 3.5, although under certain penetration rates (like 10% and 100%) random seeds have an influence on the Relative Reduced TTI, the influence of the penetration rate strongly dominates that of the random seed. Therefore, balancing the time efficiency of data collection and the accuracy of data, in all following simulations, one fixed random seed is used.

<table>
<thead>
<tr>
<th>Random seed</th>
<th>CLF</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.19</td>
<td>1.54%</td>
<td>6.86%</td>
<td>7.54%</td>
<td>8.65%</td>
<td>8.05%</td>
<td>8.11%</td>
<td>7.10%</td>
<td>5.73%</td>
<td>4.66%</td>
<td>3.26%</td>
</tr>
<tr>
<td>21</td>
<td>0.19</td>
<td>3.48%</td>
<td>7.22%</td>
<td>8.63%</td>
<td>9.01%</td>
<td>7.70%</td>
<td>7.46%</td>
<td>6.10%</td>
<td>4.06%</td>
<td>0.89%</td>
<td>-0.89%</td>
</tr>
<tr>
<td>33</td>
<td>0.19</td>
<td>5.94%</td>
<td>7.06%</td>
<td>8.84%</td>
<td>9.47%</td>
<td>9.94%</td>
<td>8.04%</td>
<td>8.65%</td>
<td>4.71%</td>
<td>4.37%</td>
<td>3.75%</td>
</tr>
<tr>
<td>64</td>
<td>0.19</td>
<td>5.16%</td>
<td>6.69%</td>
<td>9.01%</td>
<td>10.28%</td>
<td>9.95%</td>
<td>9.77%</td>
<td>8.56%</td>
<td>7.25%</td>
<td>6.87%</td>
<td>4.45%</td>
</tr>
<tr>
<td>10000</td>
<td>0.19</td>
<td>3.37%</td>
<td>6.91%</td>
<td>7.52%</td>
<td>9.39%</td>
<td>10.49%</td>
<td>8.75%</td>
<td>8.11%</td>
<td>6.85%</td>
<td>4.96%</td>
<td>3.96%</td>
</tr>
</tbody>
</table>

Table 3.5: Influence of random seeds in Network 2 (Pilot 1-3)
3.4.5.2. **Pilot 2: Influence of the penetration rate**

In this section, the influence of the penetration rate on the effectiveness of the on-trip dynamic navigation system is examined. The values of the Relative Reduced TTI under different penetration rates on different congestion levels are calculated based on simulation results in DYNASMART-P, in order to investigate how the penetration rate influences the effectiveness of the on-trip dynamic navigation system. Table 3.6 lists the range of values of situational variables in Pilot 2.

<table>
<thead>
<tr>
<th>Network density ($km/km^2$)</th>
<th>1.40 (Network 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestion level</td>
<td>CLF: 0 to 1</td>
</tr>
<tr>
<td>Penetration rate</td>
<td>10%, 20%, 30%, ..., 90%, 100%</td>
</tr>
<tr>
<td>Information update interval (minutes)</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.6: The range of values of situational variables in Pilot 2

The values of the Relative Reduced TTI based on 100 simulations is shown in Figure 3.3. Note that because when no drivers use the on-trip dynamic navigation system naturally this ITS measure makes no influences, so when the penetration rate is zero, the value of the Relative Reduced TTI is set to 0% without running any simulations. It could be found that whatever the congestion level is, **when the penetration rate is low, the penetration rate has positive relationship with the effectiveness of the on-trip dynamic navigation system, but when the penetration rate is higher than 50%, the effectiveness would decrease with the increasing penetration rate**. The result is totally consistent with **Hypothesis 3** in Section 3.3.5.
3.4.5.3. **Pilot 3: Influence of the Congestion Level**

Pilot 3 examines the influence of the congestion level on the effectiveness of the on-trip dynamic navigation system. The range of values of situational variables is shown in Table 3.7.

<table>
<thead>
<tr>
<th>Network density (km/km²)</th>
<th>1.40 (Network 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestion level</td>
<td>CLF: 0 to 1</td>
</tr>
<tr>
<td>Penetration rate</td>
<td>—</td>
</tr>
<tr>
<td>Information update interval (minutes)</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3.7: The range of values of situational variables in Pilot 3

Figure 3.3 in Section 3.4.5.2, to some extent, already shows a basic relationship between the congestion level and the effectiveness of the on-trip dynamic navigation system: with the increasing CLF, the effectiveness would be increased. But to study the relationship more precisely, the average value of the Relative Reduced TTI under different penetration rates on every congestion level is computed. Figure 3.3 is thus transformed to the following Figure 3.4.

![Figure 3.4: Influence of the congestion level on the effectiveness of the on-trip dynamic navigation system (Pilot 3)](image)

According to Figure 3.4, the average Relative Reduced TTI, in other words the **effectiveness of the on-trip dynamic navigation system would be increased with the increasing CLF**. The result is partially consistent with Hypothesis 2 in Section 3.3.5 although the hypothesized downward sloping branch is not observed. The slight difference between the observed result and the hypothesis might be due to two reasons. On one hand, Hypothesis 2 is only based on theories and deductions, in fact there are no evidences or experiments showing this assumption is really true. On the other hand, the CLF is set within a limited range in this study and thus the result cannot represent all possible conditions. Furthermore, it should be noted that the relationship shown in Figure 3.4 should be only applicable when the CLF is within 0 to 1.
3.4.5.4. **PILOT 4: INFLUENCE OF THE INFORMATION UPDATE INTERVAL**

In pilot simulation 4, the influence of the information update interval on the effectiveness of the on-trip dynamic navigation system is tested. The range of values of situational variables is shown in Table 3.8. Note that to reduce the workloads and make the result comparable, in pilot 4, all tests are based on the same congestion level (with the same value of CLF).

<table>
<thead>
<tr>
<th>Network density ($km/km^2$)</th>
<th>1.40 (Network 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestion level</td>
<td>On the same congestion level (CLF=0.80)</td>
</tr>
<tr>
<td>Penetration rate</td>
<td>10%, 20%, 30%, ..., 90%, 100%</td>
</tr>
<tr>
<td>Information update interval (minutes)</td>
<td>2, 3, 4, 5</td>
</tr>
</tbody>
</table>

Table 3.8: The range of values of situational variables in Pilot 4

![Figure 3.5: Influence of information update interval on the effectiveness of the on-trip dynamic navigation system (Pilot 4)](image)

The simulation result is shown in Figure 3.5.

According to Figure 3.5, it could be concluded that basically **the effectiveness of the on-trip dynamic navigation system would be decreased as the information update interval is increased**, which is consistent with Hypothesis 4 in Section 3.3.5.

Moreover, when the penetration rate is between 10% to 30%, the values of Relative Reduced TTI with different information update intervals are very close, and when the penetration rate is higher than 30%, the difference among the values of Relative Reduced TTI with different information update intervals becomes more and more obvious with the increasing penetration rate. This might be due to when the penetration rate is low which means only the minority of drivers use the on-trip dynamic navigation system, the decreasing information update interval can bring benefits to only a few drivers thus the influence
of information update interval is limited, but when the penetration rate is high, the majority of drivers use real-time information, the decreasing information update interval can provide better information to most drivers and thus the influence of information update interval becomes significant.

Furthermore, if information update interval is 4 or 5 minutes, the effectiveness could be negative when the penetration rate is higher than 80%, which means when most drivers use real-time information, information that is updated slowly might lead to a worse situation.

However, it should be noticed that the result shown in Figure 3.5 is a little different from Hypothesis 4 in Section 3.3.5. Because the update interval of pre-trip information and the update interval of on-trip have to be the same in DYNASMART-P, the uninformed drivers and informed drivers actually use the information with the same update interval and the only difference is that uninformed drivers cannot get the information during their trips. In other words, the on-trip dynamic navigation system with 2-minutes information update interval is in fact compared to pre-trip information with 2-minutes information update interval. Therefore, the values of Relative Reduced TTI in the four data series in Figure 3.5 are calculated based on different reference values, thus it cannot be simply concluded that reducing information update interval can improve the effectiveness of the on-trip dynamic navigation system. The accurate conclusion in fact should be the benefits from using on-trip information instead of pre-trip information would be increased with the decreasing information update interval. But to express the result simply, in the following study, such a result is still stated as the effectiveness of the on-trip dynamic navigation system would be increased with the decreasing information update interval.

3.4.5.5. PILOT 5: INFLUENCE OF THE NETWORK DENSITY

At last, Pilot 5 discusses the influence of the network density on the effectiveness of the on-trip dynamic navigation system. Three different networks are compared in this section. Network 3 (Map C.3) and Network 4 (Map C.4) are built based on Network 2 so all of these three networks have similar structure, while Network 3 is the sparsest one and Network 4 is the densest one. The range of values of situational variables is shown in Table 3.9.

<table>
<thead>
<tr>
<th>Network density (km/km²)</th>
<th>1.40 (Network 2); 1.22 (Network 3); 1.52 (Network 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congestion level</td>
<td>CLF: 0 to 1</td>
</tr>
<tr>
<td>Penetration rate</td>
<td>10%, 20%, 30%, ..., 90%, 100%</td>
</tr>
<tr>
<td>Information update interval (minutes)</td>
<td>2, 3, 4, 5</td>
</tr>
</tbody>
</table>

Table 3.9: The range of values of situational variables in Pilot 5

To ignore the influence of the information update interval, networks are compared with different information update intervals. Additionally, with every value of the information
update interval, the CLF is controlled as the same to eliminate the influence of the congestion level. Some examples are shown in Figure 3.6.

(a) 2-minutes interval, CLF=0.30  (b) 3-minutes interval, CLF=0.24
(c) 4-minutes interval, CLF=0.65  (d) 5-minutes interval, CLF=0.34

Figure 3.6: Influence of the network density on the effectiveness of the on-trip dynamic navigation system (Pilot 5)

According to Figure 3.6, the values of the Relative Reduced TTI in different networks (on the same congestion level and with the same information update interval) are indeed different but very close as well. In other words, the network density could influence the effectiveness of the on-trip dynamic navigation system but the influence is quite slight according to Pilot 5. Besides, the trend described in Hypothesis 1 in Section 3.3.5, which is that the effectiveness of the on-trip dynamic navigation system would be increased with the increasing network density, is not discovered in Figure 3.6. Therefore, how the network density affects the effectiveness still cannot be answered.

One possible reason for the limited influence might be that the network density is not the suitable variables to describe the network geometry in this specific case. Anyway, since the influence of the network density on the effectiveness of the on-trip dynamic navigation system is indeed observed although the influence is quite slight, the network density is still considered in data collection.
3.4.5.6. CONCLUSION OF PILOT SIMULATIONS

Based on the pilot simulations, it is proved that the chosen four situational variables can indeed influence the effectiveness of the on-trip dynamic navigation system, so all of them are adopted in the following data collection. Besides, the influences of the penetration rate, the congestion level and the information update interval are significant but the influence of the network density is a little bit slight.

3.4.6. DATA SETS

As mentioned above, the data is collected by simulations due to the lack of practical data. DYNASMART-P still needs available networks to conduct simulations. However, the available networks which can be used in DYNASMART-P are quite limited, so the ideal situation — choosing certain ‘large scenario’ (e.g. a country) and collecting the data on local scale (e.g. the cities in the country) is not feasible in this case. Actually, there are some available networks for DYNASMART-P, but the networks are not built based on cities in the same country or same region.

Considering the influential factors for the outputs (i.e. situational variables) can be highly controlled by the simulation tool, the relationship between the impacts of the on-trip dynamic navigation and the situational variables would not be significantly influenced by the resources of the networks. In other words, although the networks are not built based on cities in the same country or same region, the relationship between the impacts of the on-trip dynamic navigation and the situational variables, which is built based on the data collected in these networks, is still reliable.

Therefore, this study takes full advantages of the available networks to collect data for finding out the relationship between the situational variables and the impacts of the on-trip dynamic navigation system, whether the networks are in the same large scenario or not.

Due to the limited resources, networks in only four urban areas are available. To have a richer data, for each urban area, three networks are built with different values of the network density (the list of networks is shown in Table 3.10). The densest network includes primary, tertiary and secondary roads in the urban area; the medially dense network contains primary and tertiary roads in the urban area; and the sparsest network only consider primary roads. Note since the number of primary roads in Amsterdam is quite limited, so for Amsterdam, there are only two networks created — AM1 contains primary and tertiary while AM2 includes primary, tertiary and secondary roads. Highways are not included. **Networks built based on the same urban area are called homologous networks** in this
study specifically. The homologous networks have the similar networks but are distinct by the network density. Additionally, since road type is not one of situational variables, all roads in the homologous networks are set to the same speed limit and capacity to eliminate the influence of proportion of different road types.

On the other hand, although network structure is not considered as a situational variables, it is still an important characteristic of networks, hence the four cities are with different structures to show the universality of collected data. Fort Worth and Xi’an have quite typical grid network, Chengdu owns a typical radial structure while Amsterdam’s network is partially grid and partially radial.

The information of all networks is shown in Table 3.10 and the maps of cities and networks are shown in Appendix C. The range of values of situational variables except the network density is shown in Table 3.11.

<table>
<thead>
<tr>
<th>City (District)</th>
<th>Structure</th>
<th>Network</th>
<th>Network density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fort Worth, USA</td>
<td>Grid</td>
<td>FW1 (Network 3)</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FW2 (Network 2)</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FW3 (Network 4)</td>
<td>1.52</td>
</tr>
<tr>
<td>Xi’an, China</td>
<td>Grid</td>
<td>XA1 (map C.5)</td>
<td>2.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XA2 (map C.6)</td>
<td>2.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XA3 (map C.7)</td>
<td>3.10</td>
</tr>
<tr>
<td>Chengdu, China</td>
<td>Radial</td>
<td>CD1 (map C.8)</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD2 (map C.9)</td>
<td>2.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD3 (map C.9)</td>
<td>3.37</td>
</tr>
<tr>
<td>Amsterdam, NL</td>
<td>Grid + Radial</td>
<td>AM1 (map C.11)</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AM2 (map C.12)</td>
<td>3.94</td>
</tr>
</tbody>
</table>

Table 3.10: The information of networks analysed in the study

<table>
<thead>
<tr>
<th>Congestion level</th>
<th>CLF: 0 to 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penetration rate</td>
<td>10%, 20%, 30%, ... , 90%, 100%</td>
</tr>
<tr>
<td>Information update interval</td>
<td>2, 3, 4, 5 minutes</td>
</tr>
</tbody>
</table>

Table 3.11: The range of values of situational variables (except the network density) in the study

Note that to guarantee the reliability of collected data, 10 different values of CLF are chosen in each network. Therefore, 4400 data sets are collected in this study. Each data series has five elements: the network density, the information update interval, the CLF, the penetration rate and the Relative Reduced TTI.

\[11 \times 10 \times 10 \times 10 \times 4 = 4400\]
3.5. **BUILDING THE DETERMINISTIC RELATIONSHIP**

In this section, the relationship between the impacts and the situational variables is discussed and determined. Section 3.5.3 investigates the relationship existing between the impacts and the penetration rate firstly. Then the relationship between the impacts and the congestion level is generated in Section 3.5.4. After that, Section 3.5.5 discusses the relationship between the impacts and the information update interval. The influence of network density is analysed in Section 3.5.6. At last, a defined relationship between the situational variables and the impacts of the on-trip dynamic navigation system is shown in Section 3.5.7.

### 3.5.1. **Methodology**

This section explains the methods (methodologies) used in building the deterministic relationship.

#### 3.5.1.1. **The Basic Approach for Building the Deterministic Relationship**

According to Section 2.2.3, the recommended approach to build the deterministic relationship is assuming a mathematical function to represent the relationship and then conducting a regression analysis to examine if the assumed relationship is reliable and determine the parameters in the hypothetical function. However, there are four situational variables in this case, it is difficult to assume a convincing function including all situational variables at the beginning. Therefore, in this section, the deterministic relationship is built step by step through analysing the influences of situational variables separately.

In addition, to analyse the relationship between the effectiveness of the on-trip dynamic navigation system and one situational variable, the recommended approach in Section 2.2.3 is used. First of all, by observing the plotting of data series, an assumption for the relationship between the impacts and the studied situational variable, generally a function, is made. Then data collected is fitted by the corresponding regression model in MATLAB to examine if the assumed model can fit the data well. The main criterion to evaluate the goodness of fit is the value of R-squared. In addition, to discuss the relationship comprehensively, a residual analysis is made to check if the built relationship function can estimate the impacts accurately by the situational variable.
3.5.1.2. **Analysis Ordering of the Situational Variables**

The analysis ordering of the situational variables are based on the influence of each situational variable shown in the pilot simulations.

Based on Pilot 2, the relationship between the penetration rate and the impacts of the on-trip dynamic navigation system seems quite clear according to the data. The found relationship also makes sense in theory. So firstly, the deterministic relationship between the penetration rate and the Relative Reduced TTI (the effectiveness of the on-trip dynamic navigation system) is investigated. The relationship between the congestion level and the impacts is also significant — the congestion level has a approximately linear relation with the average Relative Reduced TTI, as shown in Pilot 3, so the influence of the congestion level is discussed after that of penetration rate. Besides, according to Pilot 4, the information update interval does influence the effectiveness of the on-trip dynamic navigation system but not any specific relationships (e.g. linear or quadratic) are observed. Thus this situational variable is analysed after the congestion level. Furthermore, since the influence of network density is not quite significant according to Pilot 5, it is investigated in the end.

Note that such an order is only for the convenience of data analysis and it would not influence the finally built relationship. Because all situational variables are expected to be included in the finally deterministic relationship function, additionally the evaluation of the relationship would be based on the data of all situational variables.

3.5.2. **Mathematical Definitions for Related Terms**

The mathematical definitions for the parameters, variables and terms which are used in the data analysis are made:

- \( P \) penetration rate (%)
- \( C \) CLF
- \( I \) information update interval (minutes)
- \( N \) network density
- \( Y \) the impacts of the on-trip dynamic navigation system (Relative Reduced TTI) at penetration rate \( P \) with information update interval \( I \) in a network which has a congestion level of \( C \) and network density \( N \)
- \( a_i, b_j, c_k, d_l \) parameters in the deterministic relationship
- **absolute residual**
  - \( \text{absolute residual} \)
    - the absolute value of the residual
- \( MAE \)
  - mean absolute error (Note that in this common statistical term, the error is indeed the residual, although generally the error and the residual have the different definitions in statistics. So \( MAE \) can also be interpreted as the average value of absolute values of residuals.)
### 3.5.3. Penetration Rate

#### 3.5.3.1. Assuming A Possible Relationship Based On Plots

According to Figure 3.3 in Pilot 2, the Relative Reduced TTI would increase and then decrease as the penetration rate rises. The best-known type of mathematical function which shows such a relation is the *quadratic function*. Therefore, it can be assumed that the penetration rate and the Relative Reduced TTI are related in a quadratic function. Data series thus are randomly chosen from the database and fitted using the quadratic function model in Excel. The result is shown in Figure 3.7. According to the examples, the quadratic function model indeed matches the chosen data quite well, in other words, the assumption is sensible and could be tested further. Note that since when the penetration rate is zero, the value of the Relative Reduced TTI should be zero, the constant in the quadratic functions is fixed to zero.

\[
Y_{CIN}(P) = a_1 \times P^2 + a_2 \times P = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \times \begin{bmatrix} P^2 \\ P \end{bmatrix}
\]  

\[ (3.7) \]
3.5.3.2. The Goodness of Fit

All 4400 data series collected are fitted by quadratic functions in Matlab. The corresponding fitting equations and the values of R-squared are computed. It is found 85% fits achieve a value of R-squared higher than 0.9, indicating quadratic functions can match the data very well. Therefore, if purely considering the values of R-squared, the assumption, that is the penetration rate and the Relative Reduced TTI have a quadratic function relation, is considered stand.

<table>
<thead>
<tr>
<th>Range of R-squared</th>
<th>[0.00, 0.80)</th>
<th>[0.80, 0.90)</th>
<th>[0.90, 1.00]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of all data series</td>
<td>5%</td>
<td>10%</td>
<td>85%</td>
</tr>
</tbody>
</table>

Table 3.12: Values of R-squared (fitting with Function 3.7)

3.5.3.3. Residual Analysis

Table 3.12 indeed reflects that the penetration rate and the Relative Reduced TTI have a significant quadratic function relation. But it is also found that there are 5% data series have not been well matched by quadratic functions (i.e. the value of R-squared is lower than 0.8). Therefore, to further examine if Function 3.7 can estimate the impacts accurately by the penetration rate, a residual analysis is made. Figure 3.8 shows some examples.

According to the residual analysis, it is found that the absolute residual is relative larger when congestion level is close to 0. Such a fact is also reflected via the value of R-squared. Figure 3.9 shows the distribution of R-squared value in different ranges of the CLF in an example. It can be seen that when the CLF is smaller than 0.2, the value of R-squared is often lower than 0.9, but when the CLF is higher than 0.2, the value of R-squared is often higher than 0.9. That means when the CLF is low, the quadratic function relationship between the Relative Reduced TTI and the penetration rate is not very obvious. When the network is not very congested, basically the impacts of the on-trip dynamic navigation system is quite slight, so neither the improvement of the effectiveness with the increasing penetration rate nor the drop of the effectiveness when the penetration rate is extremely high can be significant, which makes the quadratic function relationship not notable.

In addition, it is also noticed that there are certain inner trends shown in the residuals. For example, in Figure 3.8a and Figure 3.8b, the residual decreases, increases and decreases again with the increasing penetration rate, and in Figure 3.8c, the residual decreases and then increases with the increasing penetration rate. The existence of the inner trends might indicate that the natural relationship between the penetration rate and the Relative Reduced TTI is not exactly a quadratic function.
3.5. **Building the Deterministic Relationship**

Figure 3.8: Examples: residual analysis of fits using Function 3.7

(a) FW2, 2-minutes interval, CLF=1.41  
(b) FW2, 4-minutes interval, CLF=1.65  
(c) XA1, 3-minutes interval, CLF=1.84  
(d) XA1, 5-minutes interval, CLF=1.48  
(e) CD3, 3-minutes interval, CLF=0.12  
(f) AM2, 4-minutes interval, CLF=0.52
So cubic function might be adopted to make the data to be fitted better. However, although the cubic function can make the estimation of model more closer to the actual data than the quadratic function, it is very risky to use such a cubic function model. In the first place, the quadratic function model is used mainly because its mathematical feature is consistent with the relationship shown between the Relative Reduced TTI and the penetration rate: the Relative Reduced TTI would increase and then decrease with the increasing penetration rate. But the mathematical feature of cubic function is totally different. The curve of a cubic function has much more changes — it can increase, decrease then increase again, which is not consistent with the relationship between the penetration rate and the Relative Reduced TTI. So from a theoretical perspective, the quadratic function model is more suitable than the cubic polynomial function to describe the relationship between the penetration rate and the Relative Reduced TTI, especially considering the fitting result using the quadratic function model is indeed good enough. Certainly, other models like 4th or 5th polynomial function might also be used, but those models have the same drawbacks as the cubic model thus are not be considered.

In one words, the quadratic function model can represent the relationship between the penetration rate and the Relative Reduced TTI very well although there are some slight errors.
3.5. BUILDING THE DETERMINISTIC RELATIONSHIP

3.5.4. CONGESTION LEVEL

3.5.4.1. ASSUMING A POSSIBLE RELATIONSHIP BASED ON PLOTS

According to the fitting results in the previous section (an example is shown in Figure 3.10), it is noticed that the shape of fit curve using function 3.7 seems be related to the value of the CLF. According to Figure 3.10, with the increasing CLF, the fit curve becomes steeper.

![Figure 3.10: An Example of fitting data by Function 3.7 (FW2 network, 2-minutes information update interval)](image)

When the curve becomes steeper, naturally, the value of the Relative Reduced TTI under a certain penetration rate would increase, which is consistent with the finding in Pilot 3: the Relative Reduced TTI would increase when CLF increases. In addition, it is seen that the maximum value of the quadratic fitting curve would increase with the increasing CLF but the maximum value is always be reached when the penetration is around 55%.

Based on the mathematical characteristics of Function 3.7, its maximum value \(-\frac{a_2^2}{4a_1}\) is reached when the penetration rate is equal to \(-\frac{a_2}{2a_1}\). Therefore, in other words, when the CLF increases, \(-\frac{a_2}{2a_1}\) would be unchanged but \(-\frac{a_2^2}{4a_1}\) would increase. Then it is not difficult to infer that \((-\frac{a_2^2}{4a_1})/(-\frac{a_2}{2a_1})\) which is equal to \(\frac{a_2}{2}\) would increase with the rising CLF. Therefore, **the parameter \(a_2\) should have a positive relationship with the CLF.** Furthermore, since \(-\frac{a_2}{2a_1}\) would be unchanged with the increasing CLF, **the parameter \(a_1\) should have a negative relationship with the CLF.**

Therefore, considering the relationship built between the penetration rate and the Relative Reduced TTI is proved reliable, an assumption is made: **the congestion level has a deterministic relationship with the parameters \((a_1, a_2)\) in Function 3.7.** The corresponding values of \(a_1\) and \(a_2\) under different congestion levels are plotted. Some of plots are
shown in Figure 3.11 and Figure 3.12.

(a) XA2 network, 2-minutes interval  
(b) CD1 network, 3-minutes interval  
(c) AM2 network, 4-minutes interval  
(d) FW1 network, 5-minutes interval

Figure 3.11: Value of $a_1$ under different congestion levels

According to the plots, the observed relationships between $a_1$, $a_2$ and the CLF are consistent with the previous assumptions and the relationships are either quadratic or linear in most cases. Considering some of these plots show a strong quadratic function relation (for example Figure 3.11b and 3.12b) and the linear function can be included in the quadratic function, the relationship between $a_1$, $a_2$ and the CLF is assumed as Function 3.8. Furthermore, combining Function 3.7 and 3.8, Function 3.9 can be deduced.

\[
\begin{bmatrix}
  a_1 \\
  a_2
\end{bmatrix} =
\begin{bmatrix}
  b_1 & b_2 & b_3 \\
  b_4 & b_5 & b_6
\end{bmatrix} \times
\begin{bmatrix}
  C^2 \\
  C \\
  1
\end{bmatrix}
\]  

(3.8)

\[
Y_{IN}(P, C) =
\begin{bmatrix}
  b_1 & b_1 & b_3 \\
  b_4 & b_5 & b_6
\end{bmatrix} \times
\begin{bmatrix}
  C^2 \\
  C \\
  1
\end{bmatrix} \times
\begin{bmatrix}
  P^2 \\
  P
\end{bmatrix}
\]  

(3.9)
3.5. Building the Deterministic Relationship

3.5.4.2. The Goodness of Fit

Based on Function 3.9, data is fitted in MATLAB within each information update interval for each network. An example is shown in Figure 3.13. The values of R-squared in all fits are listed in Table 3.13. For each case in Table 3.13, the number of fitted data series is 100. It is found that 90.9% fits could reach a value of R-squared higher than 0.9, which means Function 3.9 can match the collected data very well. Therefore, the assumption — the congestion level has quadratic relationships with the parameters \((a_1, a_2)\) in Function 3.7 is proved.

Note that the fits in AM1 network are much worse than those in the other networks, this might be due to certain unknown flaws in the original data inputs. For example, the alternative routes in AM1 network might be so few that the on-trip dynamic navigation system cannot make significant influences, the relationship between the effectiveness and the situational variables cannot be fully reflected via the collected data. Due to the limitation of time, this issue is not analysed in details. But to build a reliable deterministic relationship, AM1 network is excluded in the following analysis.
Figure 3.13: An example of fitting data by Function 3.9 (CD3 network, 2-minutes update interval interval)

<table>
<thead>
<tr>
<th>Network</th>
<th>Information update interval (minutes)</th>
<th>R-squared</th>
<th>Network</th>
<th>Information update interval (minutes)</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW1</td>
<td>2</td>
<td>0.9296</td>
<td>CD1</td>
<td>2</td>
<td>0.9601</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9196</td>
<td></td>
<td>3</td>
<td>0.9436</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9362</td>
<td></td>
<td>4</td>
<td>0.9598</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9474</td>
<td></td>
<td>5</td>
<td>0.9522</td>
</tr>
<tr>
<td>FW2</td>
<td>2</td>
<td>0.9072</td>
<td>CD2</td>
<td>2</td>
<td>0.9527</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9192</td>
<td></td>
<td>3</td>
<td>0.9421</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9175</td>
<td></td>
<td>4</td>
<td>0.9547</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9086</td>
<td></td>
<td>5</td>
<td>0.9527</td>
</tr>
<tr>
<td>FW3</td>
<td>2</td>
<td>0.9396</td>
<td>CD3</td>
<td>2</td>
<td>0.9643</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9244</td>
<td></td>
<td>3</td>
<td>0.9534</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9345</td>
<td></td>
<td>4</td>
<td>0.9693</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9950</td>
<td></td>
<td>5</td>
<td>0.9636</td>
</tr>
<tr>
<td>XA1</td>
<td>2</td>
<td>0.9037</td>
<td>AM1</td>
<td>2</td>
<td>0.8124</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9024</td>
<td></td>
<td>3</td>
<td>0.8344</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9534</td>
<td></td>
<td>4</td>
<td>0.8289</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9666</td>
<td></td>
<td>5</td>
<td>0.8916</td>
</tr>
<tr>
<td>XA2</td>
<td>2</td>
<td>0.9244</td>
<td>AM2</td>
<td>2</td>
<td>0.9512</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9447</td>
<td></td>
<td>3</td>
<td>0.9236</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9540</td>
<td></td>
<td>4</td>
<td>0.9585</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9830</td>
<td></td>
<td>5</td>
<td>0.9431</td>
</tr>
<tr>
<td>XA3</td>
<td>2</td>
<td>0.9230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.9602</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.9836</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.9799</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.13: Values of R-squared (fitting with Function 3.9)
3.5.4.3. Residual analysis

Similarly to Section 3.5.1, to test the reliability of Function 3.9, residual analysis is done like Figure 3.14. According to the residual analysis, the difference between the estimation and the observation still exists although the value of R-squared is quite high. It is noticed that when the congestion level is low, the absolute values of residuals are small but the relative values (the proportion of the residual to the observation) are large. When the congestion level is high, the absolute values of residuals are large but the relative values are small. However, in most cases, the absolute residual is smaller than 5%. Therefore, the Relative Reduced TTI can be accurately estimated by Function 3.9.

![Figure 3.14: An Example: residual analysis of fits using Function 3.9(CD3 network, 2-minutes update interval interval)](image)

3.5.5. INFORMATION UPDATE INTERVAL

3.5.5.1. Assuming a possible relationship based on plots

It is known that the information update interval can indeed influence the effectiveness of the on-trip dynamic navigation system. Besides, the relationship found between the information update interval and the Relative Reduced TTI should be finally combined with Function 3.9. It might be a good idea to study the relationship between the information update interval and the Relative Reduced TTI by investigating how the information update interval could influence the accuracy of estimations using Function 3.9.

Therefore, firstly it is assumed that the information update interval has no influences on the accuracy of estimations using Function 3.9, in other words, cases with different in-
formation update intervals could use the same values of parameters $b_j$ in Function 3.9. Then data collected in cases with various information update interval is uniformly fitted for each network by Function 3.9 in MATLAB. The goodness of fit in each network is shown in Table 3.14. For each case in Table 3.14, the number of fitted data series is 400. According to the values of R-squared, the model CANNOT fit the data well. Therefore, the assumption made is wrong, indicating that cases with different information update intervals should use different values of parameters $b_j$ to estimate the Relative Reduced TTI. That is, when estimating the Relative Reduced TTI by Function 3.9, the values of parameters $b_j$ is related to the information update interval, because the information update interval can affect the Relative Reduced TTI.

<table>
<thead>
<tr>
<th>Network</th>
<th>FW1</th>
<th>FW2</th>
<th>FW3</th>
<th>XA1</th>
<th>XA2</th>
<th>XA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.5713</td>
<td>0.4885</td>
<td>0.5655</td>
<td>0.5162</td>
<td>0.5609</td>
<td>0.5762</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Network</th>
<th>CD1</th>
<th>CD2</th>
<th>CD3</th>
<th>AM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.4602</td>
<td>0.4661</td>
<td>0.5464</td>
<td>0.5094</td>
</tr>
</tbody>
</table>

Table 3.14: Values of R-squared (investigating the influence of the information update interval)

In addition, residuals are plotted for every network and Figure 3.15 shows some examples. Note that in Figure 3.15, the horizontal axis shows a case number which runs in a triple loop over the information update intervals (out loop), the penetration rates (middle loop) and the congestion levels (inner loop). According to Figure 3.15, it could be seen that when using universal parameters $b_j$ for cases with different information update intervals, the Relative Reduced TTI in cases with 2-minutes and 3-minutes information update interval is under estimated while the Relative Reduced TTI in cases with 4-minutes and 5-minutes information update interval is over estimated. Besides, the residuals would decrease while the absolute residuals would increase with the increasing penetration rate.

Based on the characteristics of the residuals, a bold assumption is made: the information update interval has a linear or quadratic relationship with the parameters in Function 3.9. Considering the linear function can be part of the quadratic function, the relationship is assumed as quadratic (Function 3.10). If after fitting, the values of $c_1, c_4, ..., c_{16}$ are very small, the linear relationship would be examined instead.

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ c_7 & c_8 & c_9 \\ c_{10} & c_{11} & c_{12} \\ c_{13} & c_{14} & c_{15} \\ c_{16} & c_{17} & c_{18} \end{bmatrix} \times \begin{bmatrix} l^2 \\ l \\ 1 \end{bmatrix} \quad (3.10)$$
Combining Function 3.9 and 3.10, a new relationship is represented as Function 3.11 or Function 3.12.

\[
Y_N(P, C, I) = c_1 I^2 C^2 P^2 + c_2 I C^2 P^2 + c_3 C^2 P^2 + c_4 I^2 C P^2 + c_5 I C P^2 + c_6 C P^2 +
\]
\[
c_7 I^2 P^2 + c_8 I P^2 + c_9 P^2 + c_{10} I^2 C P + c_{11} I C P^2 + c_{12} C^2 P +
\]
\[
c_{13} I^2 C P + c_{14} I C P + c_{15} C P + c_{16} I^2 P + c_{17} I P + c_{18} P
\]

(3.11)

\[
Y_N(P, C, I) = \begin{bmatrix} c_1 I^2 + c_2 I + c_3 \\ c_4 I^2 + c_5 I + c_6 \\ c_7 I^2 + c_8 I + c_9 \\ c_{10} I^2 + c_{11} I + c_{12} \\ c_{13} I^2 + c_{14} I + c_{15} \\ c_{16} I^2 + c_{17} I + c_{18} \end{bmatrix} \times \begin{bmatrix} C^2 \\ C \\ P \end{bmatrix}
\]

(3.12)
3.5.5.2. Test of assumed relationship by fitting data

To begin with, data is fitted in MATLAB for each network using Function 3.12. The goodness of fit is shown in Table 3.15. For each case in Table 3.15, the number of fitted data series is 400. According to the R-squared values, the assumed model can fit the observations quite well. Besides, the values of $c_1, c_4, ..., c_{16}$ are generally within reasonable range thus the use of quadratic relationship instead of linear relationship is sensible. Therefore, it could be proved to be true that the relationship shown in Function 3.12 is a good reflection of the relationship between the three situational variables (penetration rate, congestion level and information update interval) and the Reduced Relative TTI.

<table>
<thead>
<tr>
<th>Network</th>
<th>FW1</th>
<th>FW2</th>
<th>FW3</th>
<th>XA1</th>
<th>XA2</th>
<th>XA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.9393</td>
<td>0.9163</td>
<td>0.9135</td>
<td>0.9572</td>
<td>0.9711</td>
<td>0.9777</td>
</tr>
</tbody>
</table>

Table 3.15: Values of R-squared (fitting with Function 3.12)

3.5.5.3. Residual analysis

To further investigate if Function 3.12 can estimate the Relative Reduced TTI accurately by the situational variables, residuals are plotted for each network in Figure 3.16. According to Figure 3.16, the difference between the estimation and the observation indeed exists. Besides, it is found that the absolute residuals in cases with higher information update intervals are higher than those in cases with lower information update intervals. It means, Function 3.12 can more accurately estimate the Relative Reduced TTI when the information update interval is low than when the information update interval is high. Furthermore, the absolute residuals in cases with higher penetration rates are higher than those in cases with at lower penetration rate. It means, Function 3.12 can more accurately estimate the Relative Reduced TTI when the penetration rate is low than when penetration rate.

To investigate the residuals more comprehensively, the MAE in each network is calculated (shown in Table 3.16) and the proportion of absolute residuals in each network is represented in Figure 3.17. Note that since the Relative Reduced TTI is expressed in percentage so the MAE in Table 3.16 are shown in percentage as well.

<table>
<thead>
<tr>
<th>Network</th>
<th>FW1</th>
<th>FW2</th>
<th>FW3</th>
<th>XA1</th>
<th>XA2</th>
<th>XA3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>1.5%</td>
<td>1.6%</td>
<td>1.4%</td>
<td>2.4%</td>
<td>2.5%</td>
<td>2.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Network</th>
<th>CD1</th>
<th>CD2</th>
<th>CD3</th>
<th>AM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>2.5%</td>
<td>2.7%</td>
<td>2.5%</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

Table 3.16: MAE (fitting with Function 3.12)
3.5. BUILDING THE DETERMINISTIC RELATIONSHIP

Figure 3.16: Residuals analysis of fits using Function 3.12 (the horizontal axis shows the same case number as Figure 3.15)
3. Applying the new scaling-up method to a specific ITS application

Figure 3.17: Residuals analysis of fits using Function 3.12
According to Table 3.16, the MAE are generally around 2%, in other words, when estimating the Relative Reduced TTI by Function 3.12, the absolute error is around 2% which could be considered quite small in common sense. The accuracy of estimations using Function 3.12 is quite satisfied.

Additionally, based on Figure 3.17, the proportion of absolute residuals in every network reflects a high accuracy of estimating the Relative Reduced TTI using Function 3.12. Although some residuals are higher than 15%, 91.6% absolute residuals are smaller than 5% and 62.8% are smaller than 2%. In conclusion, it is basically accurate to estimate the Relative Reduced TTI using Function 3.12.

3.5.5.4. Further Discussion

According to the analysis in the previous section, the influence of the information update interval on the effectiveness of the on-trip dynamic navigation system could be described by Function 3.12 reliably. The information update interval thus is proved to be a significant numerical situational variable.

However, it should not be ignored that there are only four different information update intervals. So the relationship shown in Function 3.12 could be assumed true only when information update intervals are 2 minutes, 3 minutes, 4 minutes or 5 minutes. It could not be directly inferred that such a relationship is also definitely correct for other information update intervals. Furthermore, the limited number of information update intervals might also lead to over-optimistic fits. It seems easy to get a good result by regression when there are only four different values for one variable.

3.5.6. Network Density

In this section, the influence of network density on the effectiveness of the on-trip dynamic navigation system is discussed.

3.5.6.1. Primary Investigation of the Influence of the Network Density

Based on Pilot 5 (Section 3.4.5.5), it could be indeed assumed that the network density could influence the effectiveness of the on-trip dynamic navigation system since the difference of the Relative Reduced TTI in various networks is indeed observed. But on the other hand, Pilot 5 also indicates that the influence of the network density seems quite slight. Therefore, in this section, whether the network density has a significant influence on the Relative Reduced TTI and thus whether the network density is really an influential
situational variable are examined.

**Function 3.12** For now, a relationship function between the other three situational variables (penetration rate, congestion level and information update interval) and the effectiveness of the on-trip dynamic navigation system is built as Function 3.12):

\[
Y_N(P, C, I) = \begin{bmatrix} 
    c_1 I^2 + c_2 I + c_3 \\
    c_4 I^2 + c_5 I + c_6 \\
    c_7 I^2 + c_8 I + c_9 \\
    c_{10} I^2 + c_{11} I + c_{12} \\
    c_{13} I^2 + c_{14} I + c_{15} \\
    c_{16} I^2 + c_{17} I + c_{18} 
\end{bmatrix} \times \begin{bmatrix} 
    C^2 \\
    C \\
    1 
\end{bmatrix} \times \begin{bmatrix} 
    P^2 \\
    P 
\end{bmatrix}
\]

Therefore, the main goal of defining the influence of the network density on the Relative Reduced TTI is to determine if different sets of parameters \(c_k\) are needed when estimating the Relative Reduced TTI. If the network density has significant influences on the Relative Reduced TTI, using the same values of parameters \(c_k\) in cases with different network density (i.e. in different networks) would lead to inaccurate estimations.

Therefore, data collected in all networks (4400 data series) is fitted by Function 3.12 in MATLAB, resulting in a R-squared value of 0.8419, which means the model cannot fit the observations very well but the fit is also not very bad. Besides, residuals are plotted in Figure 3.18. Note that in Figure 3.18, for each network, the horizontal axis shows a case number which runs in a triple loop over the information update intervals (out loop), the penetration rates (middle loop) and the congestion levels (inner loop).

According to the figure, significant differences of residuals in various networks could be seen, for example, the residuals in XA1 and the residuals in CD1 have totally different shapes. Nonetheless, the differences of residuals among the homologous networks which have different network densities but are built based on the same urban area (e.g. FW1, FW2 and FW3) seem quite small. To further examine that, data in the homologous networks is fitted by Function 3.12 in MATLAB. Th values of R-squared is shown in Table 3.17. All values of R-squared are higher than 0.9, indicating the model can match the data very well and thus using the same values of parameters \(c_k\) can lead to accurate estimations. Considering the main difference among homologous networks is the network density, and it is proved that using the same values of parameters \(c_k\) for homologous networks can still lead to accurate estimations, the influence of the network density on the effectiveness of the on-trip dynamic navigation system is really slight.

<table>
<thead>
<tr>
<th>Network series</th>
<th>[FW1, FW2, FW3]</th>
<th>[XA1, XA2, XA3]</th>
<th>[CD1, CD2, CD3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.9079</td>
<td>0.9407</td>
<td>0.9483</td>
</tr>
</tbody>
</table>

Table 3.17: Values of R-squared (fitting with Function 3.9 for homologous networks)
Figure 3.18: Residuals analysis of fits using Function 3.12 (all networks)

**Function 3.9**  On the other hand, it should be noted that in reality sometimes a large scenario might adopt an on-trip dynamic navigation system with the uniform information update interval. If so, Function 3.9

\[
Y_{IN}(P,C) = \begin{bmatrix} b_1 & b_1 & b_3 \\ b_4 & b_5 & b_6 \end{bmatrix} \times \begin{bmatrix} C^2 \\ C \\ 1 \end{bmatrix} \times \begin{bmatrix} p^2 \\ P \end{bmatrix}
\]

would be used to estimate the effectiveness of the on-trip dynamic navigation system. Therefore, it would be better to also examine if different sets of parameters \( b_j \) should be applied in various network when using Function 3.9 to estimate the Relative Reduced TTI.

Therefore, data in cases with each information update interval were fitted using Function 3.9 in MATLAB. The goodness of fit is shown in Table 3.18 and residuals are plotted for every information update interval in Figure 3.19. Note that in Figure 3.19, for every network, the horizontal axis shows a case number which runs in a double loop over the penetration rates (outer loop) and the congestion levels (inner loop).

<table>
<thead>
<tr>
<th>Information update interval</th>
<th>2 minutes</th>
<th>3 minutes</th>
<th>4 minutes</th>
<th>5 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.8277</td>
<td>0.7636</td>
<td>0.8588</td>
<td>0.8054</td>
</tr>
</tbody>
</table>

Table 3.18: Values of R-squared (fitting with Function 3.9 for all networks)

According to the fitting result, the follows can be found:

- The model can fit the data not very well but also not very badly according to the values of R-squared, in other words, the network density is slightly influential to the effectiveness of the on-trip dynamic navigation system.
Applying the new scaling-up method to a specific ITS application

(a) 2-minutes interval

(b) 3-minutes interval

(c) 4-minutes interval

(d) 5-minutes interval

Figure 3.19: Residuals when fitting data by Function 3.9 (all networks)
• Comparing Table 3.18 and Table 3.14, the influence of network density is not that significant as the influence of information update interval to the effectiveness of the on-trip dynamic navigation system.

• Generally, the absolute residuals would increase when information update interval rises.

• Similar to the previous section, the significant differences of residuals among homologous networks cannot be seen.

**Conclusion**  Combining with the findings in the previous tests using Function 3.9 and Function 3.12, it could be concluded that network density is not significant situational variables based on current available data.

### 3.5.6.2. THE POSSIBILITY TO USE THE SAME VALUES OF PARAMETERS FOR ALL NETWORKS

Although it is concluded that the network density can only slightly influence the effectiveness of the on-trip dynamic navigation system, it is unknown whether the influence is insignificant enough so that all networks can use the same values of parameters to estimate the impacts with satisfied accuracy. Therefore, further residual analysis is made in this section to examine if all networks can use the same values of parameters.

Consistently to Section 3.5.6.1, the possibility of using the same values of parameters $c_k$ in Function 3.12 for all networks, and the possibility of using same values of parameters $b_j$ in Function 3.9 for all networks are discussed one after another.

**Function 3.12**  According to Section 3.5.6.1, the MAE for all cases when using Function 3.12 to estimate the Reduced Relative TTI is equal to 3.6%, which is basically acceptable.

Additionally, the proportion of absolute residuals is shown in Figure 3.20 — about 18% cases have an absolute residual higher than 5%. From this perspective, it is quite risky to use the same values of parameters $c_k$ for all networks.

Specifically, the proportion of absolute residuals for each information update interval is also calculated as shown in Figure 3.21. It could be found that when information update interval is equal to 2 minutes, there are even about 27% cases having an absolute residual higher than 5%.

Therefore, if the absolute residual is lower or higher than 5% can be considered as the criterion to judge whether an estimation is accurate or not, applying the same values of parameters $c_k$ for all networks when estimating the the Relative Reduced TTI cannot achieve
3. Applying the New Scaling-up Method to a Specific ITS Application

Figure 3.20: Proportion of absolute residuals in each range when fitting data by Function 3.12 (all networks, all information update intervals)

(a) 2-minutes interval
(b) 3-minutes interval
(c) 4-minutes interval
(d) 5-minutes interval

Figure 3.21: Proportion of absolute residuals in each range when fitting data by Function 3.12 (all networks)
accurate estimations. It should also be noted that if the critical value of absolute residual is expanded to 10%, it is possible to use the same values of parameters $c_k$ in Function 3.12 for all networks.

**Function 3.9** Similarly, the MAE with each information update interval when using Function 3.9 to estimate the Relative Reduced TTI is calculated and shown in Table 3.19. The proportion of absolute residuals in each range with each information update interval is plotted in Figure 3.22.

<table>
<thead>
<tr>
<th>Information update interval (minutes)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>3.2%</td>
<td>3.3%</td>
<td>3.4%</td>
<td>4.1%</td>
</tr>
</tbody>
</table>

Table 3.19: MAE (fitting with Function 3.19 for all networks)

![Graphs showing proportions of absolute residuals](image)

(a) 2-minutes interval  
(b) 3-minutes interval  
(c) 4-minutes interval  
(d) 5-minutes interval

Figure 3.22: Proportion of absolute residuals in each range when fitting data by Function 3.19 (all networks)

According to Table 3.19 and Figure 3.22, it could be seen, although the average residual with every information update interval is not high (around 3%), there are many cases (about 18%) having an absolute residual higher than 5%. Such an accuracy is not satisfied
thus using the same values of parameters $b_j$ in Function 3.9 for all networks is not desirable either. Note that it is found that the accuracy (the proportion of cases having an absolute residual lower than 5%) decreases with increasing information update interval according to Figure 3.22. When information update interval is 2 minutes, only 8% cases have an absolute residual higher than 5%, which means a high accuracy of estimation. But when information update interval is 5 minutes, about 15% cases have an absolute residual even higher than 10%, indicating a low accuracy of estimation.

Therefore, if the critical value of absolute residual is 5%, the accuracy is not high enough with any information update interval if applying the same values of parameters $b_j$ in Function 3.9 for all networks when assessing the impacts of the on-trip dynamic navigation system. If softening the critical value of absolute residual to 10%, the accuracy of estimation by using the same values of parameters $b_j$ in cases with 2-minutes, 3-minutes and 4-minutes information update interval is basically satisfied, but if information update interval is equal to 5 minutes, the estimation is not accurate enough still.

**Conclusion** In summary, if the absolute residual is lower or higher than 5% is considered as the criterion to judge one estimation is accurate or not, applying the same values of parameters $c_k$ in Function 3.12 or parameters $b_j$ in Function 3.9 for all networks is not favourable. If the critical value of absolute residual is softened to 10%, it is possible to use the same values of parameters $c_k$ in Function 3.12 for all networks or the same values of parameters $b_j$ in Function 3.9 for all networks when information update interval is 2 minutes, 3 minutes or 4 minutes.

3.5.6.3. Applying values of parameters from network to network

According the analysis above, the influence of the network density on the effectiveness of the on-trip dynamic navigation system is slight but not insignificant enough to adopt the same values of function parameters for all networks considering the accuracy of assessment (if the critical value of absolute residual is 5%).

However, it is also not desirable in reality to compute values of parameters for each single network. Because it is proved that the network density is not a significant influential factors for the effectiveness of the on-trip dynamic navigation, the feedback loop mentioned in the framework of the new scaling-up method (Figure 2.2) might be used to find another suitable situational variable. Furthermore, considering the network density shows a characteristic of network geometry, the alternative situational variable is better to be related to the network geometry as well.
According to Section 3.5.6.1, similar shapes of residuals are found in the homologous networks (Figure 3.18 and Figure 3.19). Therefore, it might be possible to use the network structure (e.g. grid and radial) as the alternative situational variables. For that matter, the same deterministic relationship would be used for homologous networks, in other words, the values of parameters in a network should be able to applied to its homologous networks.

To examine whether the network structure could be a suitable situational variable, the following analysis is made.

**Function 3.12** Tests are done to examine the goodness of applying values of parameters in Function 3.12 from a network to another network.

The result is shown in Figure 3.23 — the value in row $i$ column $j$ in the table represents the value of R-squared when applying the values of parameters resulted in network $i$ to network $j$. Note the term of *Less than 5%* means the proportion of absolute residuals lower than 5% to all absolute residuals.

![Figure 3.23](image-url)
According to the value of R-squared, it could be found that:

- When applying the values of parameters in a network to its **homologous networks**, the fits are generally good. Most values of R-squared are bigger than 0.9 although some are smaller than 0.8.

- When applying the values of parameters in a network to **networks with a similar structure**, the fits are not very good but are much better than the fits when applying to **networks with different structures**. So the network structure could be an influential situational variable.

- When applying the values of parameters in a grid network to AM2 network, the fits are better than the fits when applying the values of parameters from a radial network. This should be due to grid structures in AM2 network are more than radial structures (Map C.12).

- Generally, when applying the values of parameters in a network to its **homologous networks** or **networks with a similar structure**, 23% estimations could reach a value of R-squared higher than 0.9 and 71% estimations could reach a value of R-squared higher than 0.8.

Moreover, based on the accuracy of estimation (the proportion of absolute residuals lower than 5%), it could be stated:

- By using the values of parameters from another network, the estimations of the Relative Reduced TTI by Function 3.12 are sometimes quite accurate (green zones) and sometimes very unreliable (red zones).

- There always exists a set of values of parameters from another network which could be used to accurately estimate the Reduced Relative TTI in a network (in every column there is at least one green zone).

- Generally, the values of parameters from networks having a similar structure (including homologous networks) could lead to more accurate estimations of the Relative Reduced TTI compared to the values of parameters from a network with a different structure.

Thus, basically, when applying the parameters $c_k$ in Function 3.12 from a network to another network with the similar structure (including homologous networks), the accuracy of estimation is acceptable.
3.5. BUILDING THE DETERMINISTIC RELATIONSHIP

**Function 3.9** Tests have also been done for applying parameters in in Function 3.9 from a network to another. But since Function 3.9 does not include the influence of information update interval, such a application is done by every information update interval. An example is shown in Figure 3.24.

![Figure 3.24](image)

<table>
<thead>
<tr>
<th>Information update interval</th>
<th>Grid Network</th>
<th>Radial Network</th>
<th>Grid &amp; Radial</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-minutes</td>
<td>R-squared</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XA2</td>
<td>0.043</td>
<td>0.481</td>
<td>0.618</td>
</tr>
<tr>
<td>Less than 5%</td>
<td>54%</td>
<td>63%</td>
<td>77%</td>
</tr>
<tr>
<td>Average Residual</td>
<td>3.7%</td>
<td>3.7%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

| 3-minutes                   | R-squared    |                |               |
| XA2                         | 0.201        | 0.287          | 0.341         |
| Less than 5%                | 55%          | 43%            | 39%           |
| Average Residual            | 3.7%         | 3.7%           | 3.7%          |

| 4-minutes                   | R-squared    |                |               |
| XA2                         | 0.856        | 0.803          | 0.803         |
| Less than 5%                | 83%          | 75%            | 71%           |
| Average Residual            | 3.6%         | 3.7%           | 3.7%          |

| 5-minutes                   | R-squared    |                |               |
| XA2                         | 0.860        | 0.574          | 0.778         |
| Less than 5%                | 60%          | 60%            | 60%           |
| Average Residual            | 3.7%         | 3.7%           | 3.7%          |

According to the result, compared with Figure 3.23, the accuracy of estimation has not been improved significantly. In some cases, the accuracy becomes even lower. For example, when applying the values of parameters in Function 3.12 from XA2 network to FW3 network, 63% absolute residuals are smaller than 5%, but when applying the values of parameters in Function 3.9 from cases with 3-minutes information update interval in XA2 network to FW3 network, only 39% absolute residuals are smaller than 5%. But it is also found that when applying the values of parameters in a network to networks with a similar structure (including homologous networks), the fits are not very good but are much better than the fits when applying to networks with different structures.

**Conclusion** According to the analysis in this part, it is feasible to make an accurate estimation by using the values of parameters in Function 3.9 or Function 3.12 from a similar network. Although due to the limited number of available networks, certain networks might not find another ‘similar enough’ network to provide the values of parameters that can lead to accurate estimations, it is believed that a network can always obtain a set of reliable values of parameters from another network if there are enough available networks. Besides, the network structure seems an influential situational variable compared to the network density. But since there are only 10 networks studied, how significantly the network structure can influence the effectiveness of the on-trip dynamic navigation is still unknown. Furthermore, for now, the network structure is only divided into three types (i.e. grid, radial and grid & radial), to be a better situational variables, the classification of the network structure might be more concrete.
3.5.7. Defining the deterministic relationship

Combining all conclusions in the previous sections, the deterministic relationship between the situational variables and the impacts of the on-trip dynamic navigation can be described as follows. Note that $S$ is used to represent the network structure since the network density is not considered any more.

In cases on a certain congestion level with a certain information update interval in a certain network, the effectiveness of the on-trip dynamic navigation system could be represented by a function of the penetration rate:

$$Y_{CIS}(P) = a_1 \times P^2 + a_2 \times P = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \times \begin{bmatrix} P^2 \\ P \end{bmatrix}$$  \hspace{1cm} (3.13)

In cases with a certain information update interval in a certain network, the effectiveness of the on-trip dynamic navigation system could be represented by a function of the penetration rate and the congestion level:

$$Y_{IS}(P, C) = \begin{bmatrix} b_1 & b_1 & b_3 \\ b_4 & b_5 & b_6 \end{bmatrix} \times \begin{bmatrix} C^2 \\ C \\ 1 \end{bmatrix} \times \begin{bmatrix} P^2 \\ P \end{bmatrix}$$  \hspace{1cm} (3.14)

In cases in a certain network, the effectiveness of the on-trip dynamic navigation system could be represented by a function of the penetration rate, the congestion level and the information update interval:

$$Y_{S}(P, C, I) = \begin{bmatrix} c_1 I^2 + c_2 I + c_3 \\ c_{10} I^2 + c_{11} I + c_{12} \\ c_{13} I^2 + c_{14} I + c_{15} \\ c_{16} I^2 + c_{17} I + c_{18} \end{bmatrix} \times \begin{bmatrix} C^2 \\ C \\ 1 \end{bmatrix} \times \begin{bmatrix} P^2 \\ P \end{bmatrix}$$  \hspace{1cm} (3.15)

For cases in different networks, the effectiveness of the on-trip dynamic navigation system is influenced by the network structure. Generally, the value of parameters $b_j$ in Function 3.14 and $c_k$ in Function 3.15 in networks with the same network structure could be considered the same. Currently, there are only three types of network structure (i.e. grid, radial and grid & radial) are used to classify the networks. A more specific classification of network structure is expected in future researches.
3.6. **SCALING UP PROCEDURE FOR THE ON-TRIP DYNAMIC NAVIGATION SYSTEM**

In conclusion, based on the new scaling-up method, the scaling up procedure for assessing the impacts of the on-trip dynamic navigation system is briefly displayed in this section.

If different information update intervals are applied in the large scenario, the scaling up procedure would be like Figure 3.25. If a fixed information update interval is applied in the large scenario, the scaling up procedure would be like Figure 3.26.

---

**Figure 3.25:** Scaling up procedure when different information update intervals are applied

**Figure 3.26:** Scaling up procedure when a fixed information update interval is applied
Due to the limitation of data, the final step of aggregation cannot be conducted thus the large-scale impacts are not calculated in the case study. In the future, if the needed data is available, it would be beneficial to make a complete assessment of the large-scale impacts in a specific scenario.

3.7. CONCLUSION

Based on the analysis in this chapter, the conclusions are drawn as follows.

According to Section 3.2 and Section 3.3, the steps of determining a suitable indicator and choosing suitable situational variables in the new scaling-up method is indeed helpful.

The chosen indicator of the effectiveness of the on-trip dynamic navigation system — the Relative Reduced TTI is proved as a good indicator in the later analysis. It can describe the impacts of the on-trip dynamic navigation system quite well thus the process of finding deterministic relationship could be conducted smoothly. That means the chosen measure of traffic efficiency is very sensible, in other words, the three basic criteria for selecting the measure are also sound. Therefore, in future studies, such a method of choosing the indicator of the impacts could be used and the three criteria might be a good reference. The Travel Time Index (TTI) could be a good measure of traffic efficiency and the Relative Reduced TTI also might be a good indicator for the effectiveness of ITS in other researches.

For situational variables, although the network density is proved not influential to the effectiveness of the on-trip dynamic navigation system, the other three situational variables (penetration rate, congestion level (CLF) and information update interval) are proved very suitable as expected. Therefore, the method for choosing the situational variables and the three criteria are still reliable, which could be adopted in future studies. On the other hand, when the network density is failed to indicate its influences on the effectiveness of the on-trip dynamic navigation system, the network structure is used instead. This means the feedback loop in the framework of the new scaling-up is necessary. Furthermore, in future researches, these four situational variables (penetration rate, congestion level (CLF), information update interval and network structure) might be referential.

The recommended method for building the deterministic relationship in Section 2.2.3 is applied to this case study. The deterministic relationship is successfully built in this chapter thus this recommended method could be considered helpful.

Specifically, in the analysis, for the relationship between the situational variables and the impacts of the on-trip dynamic navigation system, there are some important findings:

- When the penetration rate is low, the effectiveness of the on-trip dynamic navigation
system has a positive relation with the penetration rate. When the penetration rate is bigger than 50%, the effectiveness of the on-trip dynamic navigation system would decrease with the increasing penetration rate.

- The effectiveness of the on-trip dynamic navigation system would increase when the network becomes more congested (CLF is within \([0, 1]\)).

- The effectiveness of the on-trip dynamic navigation system would decrease with the increasing information update interval.

- The network structure can influence the effectiveness of the on-trip dynamic navigation system. In networks with the same type of structure, the effectiveness of the on-trip dynamic navigation system is expected to be the same when the penetration rate, congestion level and information update interval are equal. However, there are only three types of network structure (i.e. grid, radial and grid & radial) are considered in this study. A more specific classification of network is expected in future researches.

- The built Function 3.13, 3.14 and 3.15 for estimating the impacts of the on-trip dynamic navigation system are reliable. They can result in accurate assessments although residuals always exist.

In summary, the possibility of using the new scaling-up method to assess the impacts of ITS is proved in this chapter. The case study also proves that the new scaling-up method is indeed a good approach.
EVALUATION OF THE NEW SCALING-UP METHOD

In this chapter, the evaluations of the new scaling-up method are made. In Section 4.1, a general evaluation of the new scaling-up method is made from a methodological perspective, which discusses the merits and flaws of the new method. In Section 4.2, an evaluation from a political perspective is conducted to point out the main contributions of the new scaling-up method to political making.

4.1. EVALUATION FROM A METHODOLOGICAL PERSPECTIVE

This section makes a methodological evaluation of the new scaling-up method, indicating its merits and flaws.

4.1.1. MERITS OF THE NEW SCALING-UP METHOD

There are many advantages of the new scaling-up method and all these merits are shown in the case study.

The most important merit (contribution) of the new scaling-up method is that it expands the application range of the scaling-up method. As concluded in Chapter 1, the current scaling-up methods are only applicable for ITS whose impacts are on the microscopic mechanisms like the lane choice, the speed, and the headway. The new scaling-up method nonetheless fills the gap because it is able to assess the impacts of ITS that make
The ability of the new scaling-up method to assess the impacts of ITS with network-wide influences is proved in the case study. According to the case study, when applied to on-trip dynamic navigation system which is a specific ITS measure making a network-wide influence, the new scaling-up method is able to result in a good assessment with acceptable accuracy — the residuals of estimations in various situations are very small.

Beyond expanding the application area of the scaling up approach, the new scaling-up method is expected to be able to improve the accuracy of scaling up by considering numerical situational variables besides categorical variables (although such an improvement is not proved in this study). Both of categorical and numerical situational variables can be used in the new method, while the current scaling-up methods are only suited with categorical situational variables. Generally, more alternate situational variables means that more characteristics of the small-scale scenarios can be featured by the situational variables, and hence the small-scale scenarios can be described more specifically and precisely. If the small-scale scenarios are presented in a more accurate way, the inputs for the scaling-up are considered more reliable, so it is expected that the accuracy of scaling up can be improved.

According to the case study, the use of numeric situational variables is proved feasible. Three numerical situational variables are taken into account in the case study: penetration rate, congestion level, and information update interval. Deterministic relationships do exist between these situational variables and the indicator of the impact. Note that the improvement in accuracy is not proved in the case study because there is no scaling-up results from the current scaling-up methods as the comparisons.

Moreover, compared to the current methods, the new scaling-up method proceeds every step in a more reasonable way. For example, the current scaling-up approaches do use an indicator to describe the impact and do adopt various situational variables when assessing the impact of an ITS measure, but the importance of choosing a suitable indicator and suitable situational variable does not attract enough attention. The reason for choosing the indicator or the situational variables is not explained either. In the new scaling-up method, there are specific steps for determining the indicator and situational variables, showing how the indicator and situational variables are chosen and why they are selected. By such steps, the choice of the indicator and situational variables is made more carefully and sensibly. Another example is related to the determination of the relationship between the impact and the situational variables. In the current scaling-up methods, the relationship is often simply assumed by the researchers, for example the impact of ITS on CO₂ emission is assumed the same on the same road type, but there is no specific evidence to show the assumption is true. So the relationship adopted in the current methods is not convincing
to some extent. In the new scaling-up method, the relationship used for scaling up is built by training data, which is more sensible compared to the simply assumed relationship in the current methods. Note that due to the relationship in the new method is built by data analysis instead of purely assumption, a feedback loop is thus included in case the chosen situational variables are found not suitable as expected during the data training.

The careful decisions of the indicator and situational variables do work in the case study. The selected indicator can represent the impact very well and most chosen situational variables are quite influential. Besides, building the relationship between the impact and the situational variables by training data also shows its advantages in the case study. For instance, in the data analysis, the expected relationship between two situational variables (network density and number of intersections) and the impact is not found, so these two situational variables are replaced by another situational variable (network structure). But if using the current methods, the assumed (expected) relationship would be used anyway as long as it is believed by the users, which sometime could be wrong.

4.1.2. Flaws of the New Scaling-up Method

However, there are still a few shortcomings of the new scaling-up method exposed via the case study.

Choosing the indicator and situational variables carefully sometimes means more time is required for this process. But this disadvantage not always exist. For example, as mentioned in Chapter 2, the indicator and situational variables can also come from the previous cases which use the new scaling method to study the same ITS measure.

In addition, massive data is needed to be trained for building a convincing relationship in the new scaling-up method. The more trained data is, the more reliable the relationship is. So the users of the new scaling-up method might need to balance the accuracy of the assessment and the amount of data that need to be collected.

The larger number of situational variables might make it more difficult to explain the relationship between the impact and the situational variables from theoretical perspective. When considering many numerical situational variables, the built deterministic relationship could be quite complicated and not linear. Under such a case, it could be very difficult to explain the relationship in a reasonable way. For example, if it is found that a situational variable has a cubic relationship with the impact, it is not easy to use a theory to interpret why it happens.
4.2. Evaluation from a Political Perspective (Political Contribution)

The core contribution of the new scaling-up method is that the various outputs (conclusions) can provide useful information to support policy making.

As mentioned in Section 1.1.2, the impact of an ITS measure plays an important role in deciding if this measure should be adopted. The impact in other words the effectiveness of the ITS can directly reflect the perceived effectiveness in the framework designed by Beuthe (Figure 1.1), and the perceived distribution of benefits and costs can be influenced by the impact of the ITS measure as well.

For example, according to the case study, the perceived effectiveness would increase with increasing congestion level. Then policy makers can know that on-trip dynamic navigation system would be more effective in more congested situations, so they might give priority to the cities with severe traffic congestion when adopting this ITS application. Besides, the case study also indicates that the effectiveness of on-trip dynamic navigation would increase with decreasing information update interval, but lower information update interval would require more advanced data collection and real-time model to provide the information in time thus cost more budgets. Based on such a perceived distribution of benefits and costs, policy makers might need to decide the balance between benefits and costs when determining the information update interval.

Apart from indicating the perceived effectiveness and the perceived distribution of benefits and costs, there are also other policy advices which could be made based on the outputs of the new scaling-up method. The built deterministic relationship(s) between a situational variables and the effectiveness can suggest what value of the related parameter should be set when adopting the ITS measure. For example, based on the case study, when penetration rate is low, the effectiveness of on-trip dynamic navigation has positive relationship with penetration rate, and when penetration rate is higher than 50%, the effectiveness would decrease with increasing penetration rate. So for policy makers, if they want to maximize the total benefits for all travellers, it might be wise to control the penetration rate around 50%. Additionally, it is found that when penetration rate is quite high, the impact of on-trip dynamic navigation could be negative. In other words, providing the same traffic information to all users can lead to a worse situation. This conclusion suggests that policy makers need to control the use of the same real-time navigation product or policy makers could ask the related companies to provide different informations to various users.

In conclusion, the outputs of the new scaling-up method can provide valuable information to policy makers.
4.3. **CONCLUSION**

The new scaling-up method is high valued from a methodological perspective. It can fill the gap in the current scaling-up approaches by expanding the application area of scaling-up methods to ITS that make network-wide influences. In addition, through using numerical situational variables, the new scaling-up method improves the accuracy of scaling up compared to the current methods. Furthermore, the new scaling-up method proceeds every step in a more reasonable way, resulting in a more reliable assessment. Apart from the merits, there are also some disadvantages of the new scaling-up method, like the possibility of more time cost and data needed, as well as the possible difficulty to explain the deterministic relationship(s) in a sensible way. On the other hand, from a political view, the various outputs (conclusions) can provide useful information to support policy making.
CONCLUSIONS AND RECOMMENDATIONS

In this chapter, firstly the conclusions are made by answering the research questions, and then the recommendations of further research/application are made.

5.1. CONCLUSIONS

In this section, the answers on the research questions are given in order to achieve the research objective:

to develop a scaling-up method for ITS that make direct network-wide influences to assess their large-scale impacts on traffic efficiency, and to evaluate the new method via applying it to a specific ITS application.

5.1.1. CURRENT SCALING-UP APPROACHES

RQ1: What are the current scaling-up methods?

There are two main scaling-up approaches in current studies. One uses the local impacts as the inputs for the large-scale model and calculates the impacts of ITS on country/EU level based on the large-scale model, called the *modelling method*. The other one can be named with the *statistical method* —— describing the local scenarios by situational variables, classifying the local scenarios into categories and calculating the large-scale impacts as the weighted average of the local impacts.

However, it is found that when assessing the impacts on traffic efficiency, the current
approaches are only applicable for ITS that have direct influences on the microscopic mechanisms like the lane choice, the speed, and the headway. A scaling-up method which is suitable for all types of ITS to assess the impacts on traffic efficiency is still missing.

5.1.2. THE NEW SCALING-UP METHOD

RQ2: What is the new scaling-up method?

To fill the gap between the current approaches and the necessary functional requirements, this study develops a new scaling-up method for ITS that make direct network-wide influences to assess their large-scale impacts on traffic efficiency.

The framework of the new scaling-up method is shown in Figure 2.2 and the graphical and mathematical interpretations are presented in Figure 2.3 and Figure 2.4. In brief, the new scaling-up method firstly chooses the suitable indicator of the impacts and situational variables, then collects needed data and build deterministic relationships between the indicator and the numerical situational variables, at last uses scaling sideways to calculate all local impacts and aggregates the local impacts to the large scale.

Compared to the current scaling-up approaches, the new scaling-up method adopts some new ideas. The new method includes the steps of determining a suitable indicator and choosing suitable situational variables to proceed the choice more in a more reasonable. The new method also uses deterministic relationships and scaling-sideways to make it possible to use numerical situational variables.

5.1.3. CASE STUDY (ON-TRIP DYNAMIC NAVIGATION)

RQ3: Is the new scaling-up method applicable for a specific ITS measure?

To provide evidence of the quality of the new scaling-up method, the new method is applied to a specific ITS application (on-trip dynamic navigation system).

Based on the framework of the new scaling up method, the deterministic relationship between the situational variables and the impacts of the on-trip dynamic navigation system is successfully built. Although the final aggregation cannot be conducted thus the large-scale impacts are not calculated due to the limitation of data, the specific scaling-up process is made for on-trip dynamic navigation is proposed. Based on this scaling process, the large-scale impacts could be easily calculated if input data is available. Therefore, the new scaling-up method is indeed applicable for this specific ITS measure. It is thus proved that the new scaling-up method is able to assess the large-scale impacts on traffic efficiency
for ITS that make direct network-wide influences.

Other findings from the case study are also valuable. For example, it is discovered that i) the Relative Reduced TTI is considered as a suitable indicator of the impacts of the on-trip dynamic navigation system. ii) the effectiveness of on-trip dynamic navigation could be significantly influenced by the penetration rate, the congestion level, the information update interval and the network structure; iii) the penetration rate has a quadratic relation to the effectiveness of on-trip dynamic navigation; iv) the effectiveness of on-trip dynamic navigation would increase when the network becomes more congested; and v) the effectiveness of on-trip dynamic navigation would decrease with the increasing information update interval. These findings can be directly adopted in other projects that study the impacts of on-trip dynamic navigation, which shows the practical contribution of this study.

5.1.4. Quality of the New Scaling-up Method

RQ4: How is the quality of the new method from a methodological perspective?

The new scaling-up method expands the application area of scaling-up methods to ITS that make network-wide influences, which achieves the research objective. The new scaling-up method also improves the accuracy of scaling up compared to the current methods by making it possible to use numerical situational variables. Besides, compared with the current scaling-up approaches, the new method can lead to more reliable assessments by proceeding every step in a more reasonable way. However, the new scaling-up method is not flawless. Using the new scaling-up might cost more time and require more data. Additionally, in the new scaling-up method, deterministic relationships are built between the impacts and the situational variables, so sometimes it is difficult to explain the relationships in a sensible way when the relationship functions are complex.

5.1.5. Political Contribution of the New Scaling-up Method

RQ5: What is the contribution of the new method to policy making?

The political contribution of the new scaling-up method is built mainly by providing useful information to support policy making. On one hand, according to the political economy model designed by Beuthe (Figure 1.1), the impacts of ITS play an important role in making policy decisions. The impacts of ITS can directly reflect the perceived effectiveness and the perceived distribution of benefits and costs. On the other hand, based on the outputs of the new scaling-up method, there are also other policy advices that could be made. For instance, the built deterministic relationship(s) can suggest what value of the related
parameters should be set when adopting ITS measures.

5.1.6. **Achievement Towards the Research Objective**

In conclusion, the designed new scaling-up method in this study is expected to be able to assess the large-scale impacts on traffic efficiency for ITS that make direct network-wide influences. In addition, by applying the new method to a specific ITS application (on-trip dynamic navigation), the new scaling-up method is proved applicable to the target ITS type. Therefore, the research objective is reached.

5.2. **Recommendations**

In this study, the new scaling-up method is only applied to one specific ITS. In the future, to further investigate the applicability of the new scaling-up method, it would be favoured to apply the new method to other ITS applications.

Moreover, although the new scaling-up method is designed to evaluate the impacts of ITS on traffic efficiency in this study, it is totally possible that the new method is also suitable for assessing the impacts on environment and safety. Future study could be done for examining the applicability of the new scaling-up method for studying the impacts on environment or safety.

There are also some specific recommendations for studying the impacts of on-trip dynamic navigation.

The data collected in the case study is from simulations, but the used simulation software (DYNASMART-P) is not a desirable tool. The update intervals of pre-trip information and on-trip information have to be set as the same in DYNASMART-P, which leads to weak control on the situational variable information update interval. The outputs of the software are not in a separate file, so users have to copy data from the interface every simulation, which costs plenty of time. Detailed explanations of the settings in the software are missing, the methodologies which are used to calculate the outcomes are thus not totally transparent. Therefore, it is recommended that a more controllable, user-friendly and transparent simulation tool should be used in further researches.

According to Section 3.4.5.1, it is found that the effectiveness of on-trip dynamic navigation can also be affected by highways and actuated signals but the influences of these two factors are not discussed in this study. Future researches can pay attention to these influences. In addition, other situational variables which are not considered in this study could also be taken into account.
There are only three types of network structure (i.e. grid, radial and grid & radial) are used to classify the networks for now. A more specific classification of network structure is expected in future researches.

Due to the limitation of data, the final large-scale impacts are not calculated in the case study. In the future, if the needed data is available, it would be beneficial to make a complete assessment of the large-scale impacts in a specific scenario.
BIBLIOGRAPHY


## Literature Related

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Reason Measure Was Not Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roadway congestion index</td>
<td>Requires use of arterial traffic data, which are not available in Smart Travel Lab. Intended for metropolitan region rather than particular corridor.</td>
</tr>
<tr>
<td>Travel rate index</td>
<td>Requires use of arterial traffic data, which are not available in Smart Travel Lab. Measure can be difficult for public to understand.</td>
</tr>
<tr>
<td>Travel time index</td>
<td>Requires use of arterial traffic data, which are not available in Smart Travel Lab. Requires separation of recurring and incident delay. Measure can be difficult for public to understand.</td>
</tr>
<tr>
<td>Travel delay</td>
<td>Requires separation of recurring and incident delay.</td>
</tr>
<tr>
<td>Buffer index</td>
<td>None. Selected for analysis.</td>
</tr>
<tr>
<td>Misery index</td>
<td>Not representative of all trips on roadway.</td>
</tr>
<tr>
<td>Travel rate</td>
<td>Included in many other calculations.</td>
</tr>
<tr>
<td>Delay rate</td>
<td>Included in many other calculations.</td>
</tr>
<tr>
<td>Total delay</td>
<td>None. Selected for analysis.</td>
</tr>
<tr>
<td>Relative delay rate</td>
<td>Measure may be difficult for public to understand because result is a number with no units.</td>
</tr>
<tr>
<td>Delay ratio</td>
<td>Measure may be difficult for public to understand because result is a number with no units.</td>
</tr>
<tr>
<td>Congested travel</td>
<td>Data required are not available in Smart Travel Lab. Formula requires length of congested roadway segment. There is no definition of what constitutes congestion on a roadway in Hampton Roads region.</td>
</tr>
<tr>
<td>Congested roadway</td>
<td>Data required are not available in Smart Travel Lab. Formula requires length of congested roadway segment. There is no definition of what constitutes congestion on a roadway in Hampton Roads region.</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Data required are not available in Smart Travel Lab. Requires information on trip objective. Most often used with transit services.</td>
</tr>
<tr>
<td>Speed reduction index</td>
<td>Measure may be difficult for public to understand because result is a number with no units. Result is relative to free flow speed, which is difficult for motorists to comprehend.</td>
</tr>
<tr>
<td>Congestion severity index</td>
<td>Requires use of arterial traffic data, which are not available in Smart Travel Lab.</td>
</tr>
<tr>
<td>Lane-mile duration index</td>
<td>Results would be poor since not all freeway segments in area collect traffic data.</td>
</tr>
<tr>
<td>Level of service</td>
<td>Is difficult to distinguish between levels of congestion once congested conditions are reached.</td>
</tr>
<tr>
<td>Queues</td>
<td>Difficult to estimate queues using available traffic data. Can be measured by use of aerial photography but is costly and site specific.</td>
</tr>
</tbody>
</table>

Figure A.1: Reviewed Performance Measures [Medley and Demetsky, 2003]
Figure A.2: Congestion survey results [Bertini, 2006]

Figure A.3: Traffic congestion measures and their suitability to assessment criteria [Aftabuzzaman, 2007]
### Figure A.4: Congestion Indices Evaluation Matrix [Rao and Rao, 2012]

<table>
<thead>
<tr>
<th>Congestion Metric</th>
<th>Simplicity</th>
<th>Ease of Data Collection</th>
<th>Stability</th>
<th>Repeatability</th>
<th>Magnitude of Congestion</th>
<th>City Comparison</th>
<th>Continuous Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Travel Time</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Delay</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>LOS and Volume</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

### Figure A.5: Common Congestion Indicators [Litman, 2013]

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
<th>Comprehensive?</th>
<th>Multi-Modal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roadway Level Of Service (LOS)</td>
<td>Congestion intensity at a particular location, rated from A (uncongested) to F (most congested).</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Multi-modal LOS</td>
<td>Congestion delays to various modes, rated from A to F.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Travel Time Index</td>
<td>The ratio of peak period to free-flow traffic speeds.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Average Traffic Speed</td>
<td>Average vehicle travel speeds at a particular location.</td>
<td>No</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Commute Duration</td>
<td>Average time per commute trip.</td>
<td>No</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Per Capita Travel Time</td>
<td>Total average time residents devote to travel.</td>
<td>Yes if for all modes</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Percent Travel Time In Congested Conditions</td>
<td>Portion of peak-period vehicle or person travel that occurs under congested conditions.</td>
<td>No</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Congestion Duration</td>
<td>Average duration of congested conditions.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Congested Lane Miles</td>
<td>Number of lane-miles congested during peak periods.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Annual Hours Of Delay</td>
<td>Hours of extra travel time due to congestion.</td>
<td>Yes if for all modes</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Annual Delay Per Capita</td>
<td>Hours of extra travel time divided by area population.</td>
<td>Yes if for all modes</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Excess Fuel Consumption</td>
<td>Total additional fuel consumption due to congestion.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Congestion Cost Per Capita</td>
<td>Hours of delay times monetized value of travel time, plus additional fuel costs, divided by area population.</td>
<td>Yes</td>
<td>Yes if for all modes</td>
</tr>
<tr>
<td>Planning Time Index</td>
<td>Earlier departure required when traveling during peak periods.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Barrier Effect</td>
<td>Walking and cycling delay caused by wider roads</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
B

DYNASMART-P FEATURES

B.1. Demand Representation

There are two possible ways to generate vehicles in DYNASMART-P. One is to use dynamic Origin-Destination (O-D) matrices to generate and load vehicles onto the network, which means an OD matrix needs to be prepared for each time period. To use this method, the number of loading intervals, a multiplication factor for demand generation (to facilitate experimentation at different demand loading levels), the starting time of each period, and the end of vehicle generation time need to be defined before running a simulation. Another method is to specify the itineraries (origin and destination) of all vehicles with or without their corresponding travel plans. This method could be done by inputting vehicle.dat and path.dat. The data in vehicle.dat would define the origin and destination for each traveller and path.dat would define a specific route for each traveller.

B.2. Traffic Control

%%%DYNASMART-P is able to model common control strategies at link junctions such as no control, yield signs, stop signs, pre-timed signals, and actuated signals. Coordinated actuated control is not explicitly modelled, but can be approximated. DYNASMART-P can also model ramp metering on freeways. Users can specify any number of ramp meters, their location, and their operational period in the network.

B.3. Different Vehicle Types

%%%DYNASMART-P recognizes four vehicle types: passenger cars (PC), trucks, high-occupancy vehicles (HOV), and buses. These vehicle types are recognized for their effect on traffic conditions (such as link capacity, speed, density, volume, etc.) and consequently path assignment. Note that PCs, trucks, and HOVs are specified as fractions of the overall vehicle fleet. In this case, the specified O-D demand matrix should reflect vehicular trips. Alternatively, the user may specify a separate O-D demand matrix to account for trucks and HOVs
in the network. In this case, there is no need to specify the fraction of trucks or HOVs in the overall vehicle fleet.

B.4. **Multiple User Classes**

Currently, DYNASMART-P handles 5 classes of users, namely:

- **Class 1 - Unresponsive**: This class of users won't respond to any type of information. Drivers in this class receive path assignments at the beginning of simulation (before trip) and then adhere to their paths throughout the entire simulation.

- **Class 2 – System Optimal (SO)**: This class of drivers follows the system optimal (SO) assignment rule. But this class is only available if the traffic assignment is iterative in the simulation.

- **Class 3 – User Equilibrium (UE)**: Drivers in this class follow the user equilibrium (UE) assignment rule. Like Class 2, this class of users is only available if the iterative consistent assignment is chosen. Besides, often, this class is used to model travellers who are familiar with the network.

- **Class 4 – En-route Info**: Drivers in this class are able to switch their routes at each intersection (although they might not) based on the real-time information. Furthermore, two criteria — indifference band and threshold bound — are used for switching decisions. The indifference band reflects a fraction of travel time improvement below which the user will not switch routes. The threshold bound reflects a time improvement (in minutes) below which the user will not switch routes. Should any of these two criteria be exceeded, the user will switch routes at the next intersection.

- **Class 5 – VMS Responsive**: This class of drivers responds to VMS information. Drivers receive path assignments at the beginning of simulation, which they adhere to unless they encounter a VMS, and possibly decide to change their paths as a result.

B.5. **Operational Modes**

DYNASMART-P could be operated in two different assignment modes. In One-Step Simulation Assignment, vehicles are assigned to current-best-paths, random paths or any predetermined paths at the departure. In Iterative Simulation Assignment, a consistent iterative assignment procedure (UE and/or SO) is applied to get the equilibrium.

B.5.1. **Mode 1: One-Step Simulation Assignment**

In Mode 1, DYNASMART-P operates as a fixed time step mesoscopic simulation assignment model. It is designed to model traffic patterns and evaluate overall network performance, possibly under real-time information systems, for a given network configuration.
(including traffic control system) and given time-dependent demand pattern. The modelling approach integrates a traffic flow simulator, a network path processing component, user behaviour rules, and information supply strategies.

In mode 1, DYNASMART-P can utilize both of two demand representation ways mentioned in B.1. In both ways, DYNASMART-P would moves vehicles until they reach their final destinations. When any vehicle is at its intermediate destination, DYNASMART-P temporarily removes this vehicle from the network so that it no longer affects prevailing network conditions. Figure B.1 illustrates the overall structure of DYNASMART-P in mode 1.

![Figure B.1: DYNASMART-P (One-step Simulation Assignment) procedure](image)

**B.5.2. MODE 2: ITERATIVE SIMULATION ASSIGNMENT**

In mode 2, DYNASMART-P allows the user to solve for an equilibrium time-dependent flow pattern in the network. The following contents would explain how this mode works.

**B.5.2.1. DEFINITION OF VARIABLES AND NOTATIONS**

- \( i = \text{origin node}, \, i \in I, \)
- \( j = \text{destination node}, \, j \in J, \)
- \( t = \text{current time interval}, \, t = 1, ..., T, \)
- \( h = \text{a travel pattern for a group of travellers at their origin, for example, travellers who} \)
have the same intermediate and final stops, preferred arrival times, and activity duration, \( h \in H \),

\[
\tau = \text{departure time interval}, \quad \tau = 1, ..., T1,
\]

\( k \) = a path in the network that starts at \( i \),

\( r_{ih}^\tau \) = number of trips with pattern \( h \) generated at \( i \) during departure time interval \( \tau \),

\( r_{ijk}^\tau \) = number of travellers who depart from \( i \) to \( j \) assigned to departure time interval \( \tau \) and path \( k \),

\( y_{ijk}^\tau \) = auxiliary number of travellers who depart from \( i \) to \( j \) assigned to departure time interval \( \tau \) and path \( k \), (number of travellers assigned to path \( k \) based on all-or-nothing assignment),

\( T_{ta}^\tau \) = travel time on link \( a \) at the beginning of period \( t \), and

\( x_{ta}^\tau \) = total number of travellers on link \( a \) at the beginning of period \( t \).

\[ B.5.2.2. \text{Problem Statement} \]

Given the number of travellers that have the same travel pattern \( h \) at each origin \( i \): \( r_{ih}^\tau \forall i \in I \), \( \forall h \in H \) and \( \forall \tau \), the objective is to determine a time-dependent assignment of vehicles to the different network paths so as to minimize the travel time for each individual traveller.

\[ B.5.2.3. \text{Solution Algorithm} \]

The steps of solution algorithm are explained as follows.

- **Step 0**: Set the iteration counter \( t = 0 \). Assign the activity-based demand, \( r_{ih}^\tau \forall i, \tau \) and \( h \) to initial set of feasible paths \( k \in k_{ij} \), where \( j \) is the first destination in the travel plan \( h \). Then the initial solution is given by \( r_{ijk}^{T,0} \forall i, h, \tau, \) and \( k \).

- **Step 1**: Under the set of departure time and path assignments \( r_{ijk}^{T,0} \), perform traffic network simulation to get the corresponding link travel times: \( T_{ta}^\tau, \forall t, a \).

- **Step 2**: For each departure time interval \( \tau \), compute the set of shortest travel time paths between each origin-destination pair.

- **Step 3**: Perform all or nothing assignment for all travellers \( r_{ij}^{T,t} \). This gives an auxiliary number of vehicles on paths for each departure time interval \( y_{ijk}^{T,t} \forall i, j \), \( j \) and \( \tau \).

- **Step 4**: Update the path by checking if \( k^* \in k_{ij} \) and include it if it does not, \( \forall i \) and \( h \). Assignments for the next iteration \( r_{ijk}^{T,t+1} \) are obtained using the method of successive averages, \( \forall i, h, \tau, \) and \( k \):

\[
r_{ijk}^{T,t+1} = \frac{1}{(t+1)} \cdot [y_{ijk}^{T,t}] + \left(1 - \frac{1}{(t+1)}\right) \cdot [r_{ijk}^{T,t}] \tag{B.1}
\]
• **Step 5:** Check the convergence criterion that is based on the difference in numbers of vehicles assigned to various departure time intervals and paths over two successive iterations. Hence, assignment to the next iterations \( r_{i \theta k}^{t,i+1} \) are compared with current path assignments \( r_{ijk}^{t,i} \), \( \forall i, j, \theta \) and \( k \):

\[
\left| r_{ijk}^{t,i+1} - r_{ijk}^{t,i} \right| \leq \varepsilon , \text{where } \varepsilon \text{ is a predefined threshold} \tag{B.2}
\]

• **Step 6:** The number of cases, \( N(\varepsilon) \), in which the above absolute value is greater than \( \varepsilon \) is recorded.

• **Step 7:** Specify a pre-set upper bound, \( \Omega \), on the number of violations, \( N(\varepsilon) \), terminate the algorithm if the number \( N(\varepsilon) \leq \Omega \), and output the joint departure time-path assignments \( r_{ijk}^{t,i} \) as the solution to the assignment problem. On the other hand, if \( N(\varepsilon) > \Omega \), the convergence criterion is not satisfied. Update the iteration counter \( \partial = \partial + 1 \) and go to step 1 with the new path assignments \( r_{ijk}^{t,i+1} \).
C

MAPS OF NETWORKS
Figure C.1: Map of Network 1
Figure C.2: Map of Network 2 (FW2 network)
Figure C.3: Map of Network 3 (FW1 network)
Figure C.4: Map of Network 4 (FW3 network)
Figure C.5: Map of XA1 network
Figure C.6: Map of XA2 network
Figure C.7: Map of XA3 network
Figure C.8: Map of CD1 network
Figure C.9: Map of CD2 network
Figure C.10: Map of CD3 network
Figure C.12: Map of AM2 network