Monitoring bicycle volumes and flows in Enschede

PDEng assignment to provide insight in bicycle volumes and flows in the municipality of Enschede for bicycle policy making

Sander Veenstra

September 24, 2015

UNIVERSITY OF TWENTE.
Monitoring bicycle volumes and flows in Enschede

providing insight in bicycle volumes and flows in the municipality of Enschede for bicycle policy making

Sander Veenstra MSc
s0023264

University of Twente
Faculty of Engineering Technology (CTW)
Centre for Transport Studies (CTS)

PDEng assignment
Final report

Supervisors:
Prof. Dr. Ing. K.T. Geurs
Dr. Ir. T. Thomas

External supervisors (gemeente Enschede):
G. Spaan
R.J. van den Hof

September 24, 2015
Preface
Voor u ligt het eindrapport van het ontwerpgedeelte van mijn Professional Doctorate in Engineering (PDEng). Een PDEng programma is een post-doctorale ontwerpopleiding waarin een trainee zich richt op een concreet ontwerpprobleem in de civiele wereld. In het kader van dit PDEng programma heb ik mij gericht op het ontwerpen van een monitoringstool voor fietsstromen in Enschede. Dit rapport is een beschrijving van enkele jaren aan onderzoek naar de potentie van de aanwezige databronnen, het verwerken van de data tot relevante informatie en het presenteren van de verkregen informatie. Het idee dat er met de data uit verkeerslichten meer te doen is dan alleen het schatten van intensiteiten en wachtten voor auto’s, en dat het ook een bron kan zijn voor fietsintensiteiten, heeft geleid tot project waar verschillende onderdelen van de verkeerskunde bijeen zijn gekomen: het verzamelen van verkeersdata, het verwerken tot verkeersinformatie en het ontsluiten van die informatie in een applicatie. Maar ook het klassieke 4-stapsmodel: ritgeneratie, distributie, modal split en toedeling zijn onderdeel van dit project. Met name het ontsluiten van de verkeersinformatie heeft me tijdens het proces veel voldoening gegeven. Ik heb mezelf nieuwe kennis eigen gemaakt op het gebied van databases en het maken van een applicatie. Deze kennis heb ik meteen kunnen toepassen: learning by doing. Met gepaste trots presenteert dan de applicatie waarin de fietsdata binnen de gemeente Enschede is samengebracht en inzicht biedt in de fietsstromen in Enschede. Dit rapport beschrijft de verschillende stappen van het verzamelen van data, via het verwerken tot informatie naar het ontwikkelen van een applicatie die inzicht biedt in de fietsstromen in Enschede.
Summary

The importance of cycling in urban areas is increasing, as sustainable modes of transport are the focus of urban transport policy nowadays. However, there is hardly any information available concerning bicycle traffic with the road authorities. Where road authorities often have monitoring strategies for car traffic, monitoring bicycle traffic is often more complicated. Cyclists are less bound to major arterials and travel shorter distances and are therefore more difficult to capture in traffic data. As a result, local transport models typically are not very well suited to model bicycle trips and cannot be used to offer quantitative support and justification to bicycle policies. The lack of data on bicycle traffic hampers municipalities to plan and improve bicycle facilities. On the other hand, in the traffic network local road authorities already collect traffic data, but this data isn’t exploited to its full extent. At signalised intersections data is collected with inductive loops for local traffic light control which potentially provides a continuous source of traffic volumes at multiple locations in the traffic network, however this data is not used in the transport policy process. In combination with the National Travel Survey (Onderzoek Verplaatsingen in Nederland), a nationwide travel survey with data about trips and their characteristics, it can be an extensive source of bicycle information (i.e. bicycle volumes at signalized intersections, travel behaviour and bicycle flows on the network) to be used in bicycle monitoring and evaluation schemes.

In an effort to address the opportunities in bicycle monitoring and to assist in the development of a more data-supported urban transport policy making process the objective of this project is to develop a monitoring tool to be used in the development, monitoring and evaluation of urban bicycle policies. There are three basic elements concerned with the monitoring tool:

- Use and disclose currently available bicycle data sources
- Combine data sources to get enriched information about bicycle traffic flows
- Provide a tool to present the information and to enable analysis of the information

The data sources used in this project consist of traffic counts at signalized intersections and trip data from the NTS.

Traffic light data and bicycle volumes

In Enschede approximately 50 signalized intersections are equipped with a control system that collects data from all inductive loops and signal groups. A substantial fraction of the inductive loops are available on (separate) cycle paths. A comparison of visual counts with inductive loop detections showed that inductive loop data at signalized intersections can be used to accurately estimate bicycle volumes (Veenstra et al., 2013). Around and below 50 counts per 15 minutes, inductive loop counts correspond well with visual counts. For higher volumes, the inductive loop counts underestimate actual bicycle volumes. Applying a correction factor (correcting for indistinguishably small headways at higher volumes) result in an accurate estimation of bicycle volumes also at higher volumes. Traffic light data provide an extensive and valuable source of data concerning bicycle volumes as it is a continuous, widely available and low-cost data source, that enables the monitoring of the dynamics of traffic volumes.

National Travel Survey and travel behaviour

Many municipalities in The Netherlands use the data from the National Travel Survey in an effort to get an understanding of the travel behaviour (trip generation, trip length distribution and modal
split) in their respective municipality. The data can offer an aggregated overview of the travel behaviour (on a municipal level) and can show the evolution of travel behaviour patterns over time. However, the number of respondents per municipality is too limited to extract information about bicycle flows or to assist in pinpointing bicycle flow issues or potential network developments on a municipal level.

**Bicycle flow estimation**
Combining the data from the NTS with the traffic light data gives an insight in urban bicycle flows for the municipality of Enschede. From the NTS the trips in and around Enschede were selected. The NTS editions from 2004 to 2013 were stacked to acquire a reasonable mass of trip data (i.e. 8216 bicycle trips in and around Enschede). Aggregates such as average daily bike trip rate per inhabitant and total trip generation of the various postal zones in Enschede were extracted to get an overview of the trip production, trip attraction and distribution of bicycle traffic in Enschede. An origin-destination matrix extracted from the NTS was assigned to a traffic network according to the All-Or-Nothing procedure on the shortest path in distance. The resulting bicycle loads were then compared with the actual counts at the signalized intersections and a matrix calibration procedure was conducted to align the matrix with the counts.

This study implies a new information source for urban bicycle traffic can be generated by combining trip data from the NTS and bicycle volumes from traffic lights in the network. Although some assumptions in the bicycle flow estimation process may not hold at all time (e.g. cyclists do not always choose the shortest path), the estimation of bicycle flows provides a very relevant and valuable information source for an urban bicycle monitoring scheme. The information about bicycle flows can support road authorities in developing more effective bicycle policies and help in evaluating specific bicycle measures on a network scale (e.g. comparing bicycle volumes before, during and after a bicycle measure) and municipalities can monitor the use of the bicycle in the urban transport system over the years.

**Database and application design**
The bicycle related traffic information is stored in a database. In the design and implementation a pragmatic approach was chosen. The database can be considered as ‘static’. On a yearly basis new data can be added. On top of the database an application is designed to handle the requests of the user and to present the information in a comprehensible fashion.

The main functionalities of the application are derived from discussions with the municipality of Enschede as primary user. For the municipality the main functionalities are listed:

- Central database for bicycle data (traffic lights, NTS, manual counts)
- Control, maintain and update the database
- Overview of bicycle volumes on main bicycle arterials
- Volumes and their dynamics of specific arterials
- Overview of travel behaviour in Enschede
- Overview of estimated bicycle flows (origins and destinations)

The resulting application will entail four main functionalities to fulfil the requirements of the end-users of the application:
1. Traffic counts and their resulting traffic profiles
2. Traffic flows and the resulting traffic loads on the network
3. Traffic generation and modal split of neighbourhoods
4. Manage and update data sources and information

These four functionalities were addressed in the design of the application. The following figures show the tabs in the application addressing the four functionalities.

Figure 0-1: final design of application

The first tab (upper left) shows the bicycle volumes at the signalised intersections. The user can select the count location of interest and show the daily volume profile. The selected count location is presented on a map for spatial reference. The second tab (upper right) is concerned with the travel behaviour and OD-relations from the NTS. The user can select the municipality or area of interest and show the trip generation, trip length distribution or modal split. At the same time the OD-relations of the selected area are presented on a map. The third tab (lower left) deals with the estimation of bicycle flows. Firstly, the user can create an OD-matrix. Secondly, the user can decide on how to assign the OD-matrix to the network. The results are presented as loads on links on a map. The fourth tab (lower right) is used for data management. New editions of the NTS and traffic light data can be loaded into the database here.

Directions for further research
This bicycle monitoring tool provides insight in the bicycle volumes and flows in the municipality of Enschede. As it is a demo-version there are components of the tool that may require further research and some assumptions may require a more in-depth investigation. The method of converting detections at inductive loops is only dependant of the bicycle volume itself. The configuration and
placement of the loop detector may also have influence, but was not taken into account in this study. The bicycle route choice component assumes all cyclists choose the shortest route (distance). The recent developments in automatically detecting and registering trips via a dedicated application on smartphones may initiate studies to clarify bicycle route choice behaviour, as this new data source holds the actual routes cyclists take. In addition a more extensive bicycle network is needed. The currently used network is rather fine-grained and contains a major cycling paths but lacks further characteristics that may influence route choice, such as comfort and obstacles.

The database and the application are currently designed to present the bicycle information according to the requested analysis of the user and is one-directional. A future version may allow contributions of users. This may open new possibilities for the municipality as road authority to acquire qualitative information from the general public about the bicycle traffic system. Moreover, the monitoring tool is currently designed as a descriptive tool presenting bicycle related information from previous years, but can be extended to a predictive tool for modelling future bicycle flows. Allowing the user to add bicycle links and to make demographic forecasts will enable the tool to predict future bicycle flows. This allows the municipality to evaluate the effects of bicycle policy measures on forehand.
# Table of Contents

Preface ......................................................................................................................... 3

Summary ......................................................................................................................... 4

Traffic light data and bicycle volumes ........................................................................... 4

National Travel Survey and travel behaviour ................................................................... 4

Bicycle flow estimation .................................................................................................. 5

Database and application design ..................................................................................... 5

Directions for further research ...................................................................................... 6

Table of Contents .......................................................................................................... 8

1 Introduction .................................................................................................................. 10

1.1 Problem definition .................................................................................................. 10

1.2 Objective ................................................................................................................ 11

1.3 Approach ................................................................................................................. 12

1.4 Outline ..................................................................................................................... 13

2 Data sources ............................................................................................................... 16

2.1 Traffic light data ..................................................................................................... 16

2.2 National Travel Survey .......................................................................................... 17

2.3 Traffic network ....................................................................................................... 18

2.4 Other data sources ................................................................................................ 20

3 Bicycle volumes ......................................................................................................... 22

3.1 Data sources .......................................................................................................... 22

3.2 Measurement biases in using inductive loop detectors .......................................... 24

3.3 Comparison of visual and inductive loop counts .................................................... 26

3.4 Discussion .............................................................................................................. 28

4 Bicycle flow estimation ............................................................................................. 29

4.1 Data sources .......................................................................................................... 29

4.2 Traffic network selection ....................................................................................... 34

4.3 Method .................................................................................................................... 35

4.4 Results .................................................................................................................... 41

4.5 Discussion .............................................................................................................. 48

5 Database design ......................................................................................................... 52

5.1 Introduction ............................................................................................................ 52

5.2 System definition .................................................................................................. 52

5.3 Data collection and analysis ................................................................................... 53

5.4 Database design .................................................................................................... 58
1 Introduction

Traffic information for car drivers is a well-established information source. While travelling a car driver can acquire the latest news about congestion and road works via the radio or navigation system. This implies a monitoring strategy is present that collects and distributes this vehicle related traffic information. Although the bicycle becomes an increasingly more important mode of transportation (especially in the urban context) there is hardly any information available concerning bicycle traffic with the road authorities. This chapter describes the motivation, the objective and the approach of the project to provide a means for gaining insight in urban bicycle volumes and flows in Enschede.

1.1 Problem definition

The importance of cycling in urban areas is increasing, as sustainable modes of transport are the focus of urban transport policy nowadays. In the Netherlands, more than 25 per cent trips are made by bicycle. The bicycle is even more important in urban transport. In various provinces and municipalities the use of the bike is stimulated. These governmental bodies commit to improve bicycle infrastructure, improve bicycle safety and upgrade bicycle facilities such as parking (Regio Utrecht, 2013, Provincie Groningen, 2012, Stadsregio Amsterdam, 2015). For example, in Amsterdam, the share of bicycle in trips made within the city of Amsterdam has increased from 33% in 1986 up to 47% in 2008. The car and public transport have a share of 31% and 22% in Amsterdam’s urban transport. Although the grow in bicycle use is widely seen as a positive development, in some areas the increase in bicycle use caters for issues with bicycle congestion and throughput (Regio Utrecht, 2013). Local governments in the Netherlands, who are responsible for bicycle planning, often have well-established procedures for collecting, summarizing, and disseminating motor vehicle traffic volumes, but these procedures do not generally include system wide bicycle volume data. Cyclists are less bound to major arterials in comparison with cars and travel shorter distances and are therefore more difficult to capture in traffic data. Only recently municipalities in The Netherlands initiate programmes to structure more network-wide monitoring scheme for cyclists. Municipalities typically use data from national travel surveys, combined with visual counts. The lack of data on bicycle volumes hampers municipalities to plan and improve bicycle facilities. As a result, local transport models typically are not very well suited to model bicycle trips.

On the other hand, in the traffic network local road authorities collect a lot of bicycle data, but this data isn’t exploited to its full extent. The main example within the municipality of Enschede is the signalised intersections. The control systems of signalized intersections in Enschede rely on the detections of multiple inductive loops near the traffic lights collecting data on arrivals of vehicles and bicycles at the respective intersection. In theory this provides an extensive source of traffic data at the main arterials of the urban transport system for both car traffic and cyclists; however it is not used in this way. The data is only used within the traffic light control system itself to regulate and optimize the traffic flows at the intersection locally. Moreover, the available data sources are generally used separately. For example, the developments in the transport system in Enschede are extracted from the NTS because it offers an overview of trip rates and modal split. Other sources could add to this overview and shed a different light on the matter (e.g. traffic volumes at certain locations in the network). Using the information from different data sources may provide a more reliable picture of the issue at hand as it is viewed from various perspectives. Moreover combining data sources may have the potential of creating synergy in uniting the advantages and specific
characteristics of the different sources and providing a richer data source answering to a larger field of transport policy relevant issues.

The problem in urban transport policy nowadays is that the shift of focus from the car towards stimulating the use of the bike is not supported by proper monitoring schemes to enable the monitoring and analysis of the entire urban transport system. The current efforts in collecting traffic and transport data can be better exploited by disclosing and using the currently available data sources and by combining the data sources where possible. This can potentially cater for synergy in terms of richer data sources that combine the strengths of the respective data sources. Processing the data into policy relevant bicycle traffic information and presenting it in a comprehensible tool will then enable the use of data from the local traffic network in urban transport policy making in which the bicycle can take a more prominent place.

1.2 Objective
In an effort to address the opportunities in bicycle monitoring and to assist in the development of a more data-supported urban transport policy making process the objective of this project is to develop a monitoring tool to be used in the development, monitoring and evaluation of urban bicycle policies. There are three basic elements concerned with the monitoring tool:

- Use and disclose currently available bicycle data sources
- Combine data sources to get enriched information about bicycle traffic flows
- Provide a tool to present the information and to enable analysis of the information

Currently available data consist of traffic counts at signalized intersections and trip data from the National Travel Survey (NTS). Besides for traffic light control optimization the traffic counts can be used to monitor traffic volumes on a network-wide scale. Because this data source is continuous in principal the dynamics can be investigated and trends can be observed. For example, the developments of traffic congestion can be monitored before during and after traffic measures have been implemented. Comparisons of traffic volumes between days, months and years can be constructed and can be an important tool in the urban transport policy.

The trip data from the NTS is used to get a general overview of the travel behaviour of Dutch citizens. This data source excels in providing a representative overview of trip related characteristics for The Netherlands such as trip generation, trip length distribution and modal split. For a municipality on its own the number of cases in this survey every year is only sufficient to get a aggregated overview of for example the modal split. Especially when trying to get disaggregate information such as an origin-destination matrix and trip generation of individual neighbourhoods the NTS lacks the critical mass.

Combining the traffic counts with the travel behaviour data may unite the advantage of a large source of continuous traffic counts with specific trip characteristics of travel behaviour to get a better overview of traffic flows in the urban transport system. This new information may assist the road authority in improving the urban transport network. For example is investigate the bottlenecks in the network and how best to resolve them.

The focus of this research lies on the bicycle. Although the constructing of a traffic monitoring strategy for all transport modes would provide a better insight in the urban transport system as a whole, the bicycle is receiving increasingly more attention in transport policy making the monitoring
of this modality is still inadequate and needs exploration. Enschede will serve as a case study to disclose and combine the data sources into policy relevant bicycle information.

1.3 Approach

This section describes what data sources are available for the case of Enschede and how to disclose, use, combine and present them to employ the efforts currently conducted in traffic monitoring and to assist in policy making and evaluation of the urban transport system.

Basically the core of the system under design is a database with input from the available bicycle data sources. In the database the data is filtered and processed into relevant bicycle information. On top of the database an application is develop for the end-user to access the bicycle information within a policy relevant context.

![Diagram of the system](image)

**Figure 1-1: project overview**

There are two main disciplines involved in the project: (1) transport engineering and (2) computer science. In the transport component the various bicycle related data sources are filtered and processed into bicycle policy relevant information. The information is stored in a database that also includes a traffic network to hold the bicycle traffic information. The information is presented via an application allowing the user (i.e. primarily the municipality as road authority) to view the information and conduct analyses for evaluation purposes. The system is influenced by the urban transport policy cycle in the sense that the system contributes to the monitoring and evaluation. This is affected by the access to data sources, requested information as input for the urban transport policy cycle and transport policy measures in the urban network.

Going more into detail the project can be broken down according to the following figure. It effectuates the structure as presented in Figure 1-1 in accordance with the case of Enschede and defines the system under design.
The starting point of the system under design is the currently available data sources in Enschede (trip data from the NTS and traffic volumes from the traffic lights). Together with the bicycle network, providing the reference to the transport network, the data is filtered, processed and combined within a dedicated database to contain bicycle traffic information of Enschede. The database needs to provide access to an external service, in this case an application, to disclose the bicycle traffic information. The application, as the second major component, provides a structure for the information to be presented in a convenient and comprehensible way. It requires an appropriate structure to provide the requested functionalities to the end-users and spatial reference to maps of the area. The combination of the two components will provide insight in the bicycle traffic information currently present within the municipality of Enschede.

1.4 Outline
The report is divided in accordance with the disciplines involved in the project. The first part the transport engineering component of the project is highlighted containing a description of the data sources and the data processing into policy relevant bicycle information. The second part presents the computer science component and describes the design of the tool (i.e. the database and the application) to present the bicycle information as mentioned in part one.

In chapter 2 the available data sources are briefly described. The process of retrieving data at signalized intersections, the scheme and procedure of the National Travel Survey and data sources that may be available in the future are mentioned. Moreover, the required traffic network as the carrier of the bicycle traffic is introduced. In the third chapter the traffic light data is described in more depth and a method is developed to process the data into estimations of bicycle volumes at the intersections. In the fourth chapter this bicycle volume information is then combined with the trip data from the NTS in an effort to construct bicycle traffic flows on the network of Enschede. This concludes the transport engineering part of the report. In the second part of the report the computer science component of the design is presented. Chapter 5 describes the structure of the database.
containing the bicycle information generated in the first part. Chapter 6 shows the development of
the application and the final result is presented. Chapter 7 discusses directions for further research.
Part I: Traffic engineering

The first of two parts of this report is devoted to the traffic engineering component of the project. It consists of the gathering bicycle data and the processing and combining the data into policy relevant bicycle information, whereas the second part describes the operational structure and the visualization of the bicycle volumes and flows for the actual storing and presentation of the bicycle information. In part I the available data sources for cycling in Enschede are introduced. The municipality of Enschede, like other Dutch municipalities, has an extensive network of traffic sensors for detecting traffic at signalized intersections. Traffic volumes (including for the bike) can be extracted from these sensors (i.e. via inductive loop detectors). A second main data source is the National Travel Survey, containing trips conducted in The Netherlands. This source can provide an aggregate overview of travel behaviour and traffic flows can be extracted based on the origin-destination data in the NTS. In chapter 2 these data sources are introduced in further detail.

The data from traffic lights needs to be processed to produce information about traffic volumes, because detections on a bicycle path do not directly represent the number of cyclists passing that detector. In the third chapter a method is developed to process the data from inductive loops at bicycle paths near traffic lights into estimations of bicycle volumes.

In the fourth chapter the data from the NTS is combined with the data from the traffic lights in order to construct an estimate of the bicycle flows in Enschede. The NTS provides a small sample of the origin-destination relations in Enschede, while the traffic light data represent the actual traffic loads. The combination of these data sources may enable the modelling of bicycle flows through the city of Enschede. The fourth chapter describes the study of combining the data sources in an effort to construct urban bicycle flows in Enschede.
2 Data sources

Especially in the urban environment the bicycle has the main focus of road authorities in terms of transport policy making as it is a sustainable and non-polluting mode of transport. Although cycling has a large share in the Dutch transport system, the knowledge about bicycle volumes and flows is by far not as extensive as for the car. Municipalities do invest in traffic monitoring, but the efforts are not exploited to the fullest especially for the bicycle. In this chapter the available data sources are discussed and their opportunities in generating policy relevant bicycle information.

2.1 Traffic light data

In the Netherlands many signalized intersections have a vehicle actuated traffic light control system. In Enschede approximately 50 signalized intersections are equipped with a control system that collects data from all inductive loops and signal groups and puts these in a central database. Most inductive loops are used for car traffic. However, a substantial fraction of the inductive loops are available on (separate) cycle paths. For these cycle paths, a central database (located at the manufacturer of the traffic light control system) stores the timestamps of bicycle detections.

Apart from the operation of the traffic signal control, the data of the inductive loops can be used to estimate volumes at signalized intersections (Nordback and Janson, 2010, Kidarsa et al., 2006, Dharmaraju et al., 2001). A typical signalized intersection equipped with inductive loops is shown in Figure 2-1.

![Figure 2-1: Inductive loop configuration at a signalized intersection](image)

The figure shows that the signal group for cyclists uses two inductive loops to detect incoming cyclists. Cyclists pass a distant loop first and stop just before the stop line. Near the stop line there is
a second loop. Normally a cyclist occupies the stop line loop while waiting for the traffic light. The loops generally are parallelogram-shaped (approximately 1.50 meter wide and 1 meter long). The distant loop is located approximately 15 meters prior to the stop line. The configuration as described before is used at most signalized intersections in Enschede.

The raw data from the inductive loop detectors consist of ‘events’ with the associated time stamps indicating whether or not a particular detector is occupied by a vehicle (i.e. a bicycle in this case). When the disturbance of the magnetic field induced by the inductive loop exceeds a particular threshold an event is triggered, indicating the loop is occupied. Also when the magnetic field is restored to normal (i.e. the cyclist leaves the loop) another event is triggered stating that the loop is unoccupied. The data imply that traffic volumes can be extracted by aggregating and counting the occupancy of the inductive loops. In the case of this research 15-minute aggregates are used for all detectors. We used these 15-minute aggregates as the basic traffic light data.

Traffic light data can be an extensive and valuable source of data concerning bicycle volumes as it is a continuous, widely available and low-cost data source, that enables the monitoring of the dynamics of traffic volumes.

### 2.2 National Travel Survey

Many municipalities in The Netherlands use the data from the National Travel Survey in an effort to get an understanding of the travel behaviour in their respective municipality. The data can offer an aggregated overview of the travel behaviour in terms of trip generation, trip length distribution and modal split and can show the evolution of these variables over time. However, as it is a nationwide survey, the number of respondents per municipality is too low to extract information about bicycle flows on a daily or even a yearly basis.

In the Netherlands a national travel survey is conducted on a yearly basis. Since 1978 Statistics Netherlands and since 2004 the Ministry of Infrastructure and the Environment (Rijkswaterstaat) investigates the mobility of Dutch citizens. The objective of the survey is to provide the ministry with information about the daily mobility of the respondents for policy making and research purposes. Respondents are asked to report on one specific day. Moreover, personal and household characteristics are reported. The respondents in the survey are representative for the Dutch population using a weight factor calculated after the survey.

Respondents are selected and receive a letter with instructions to fill in an internet-based questionnaire to report on the trips they make on an indicated day. If the respondents do not reply on the call the fill in the online questionnaire, the respondents are approached by phone to conduct the survey. Otherwise the respondents are approached to conduct the survey face-to-face. The survey was based on all members of households from 2004 to 2009. From 2010 the respondents are approached based on personal characteristics and their location of residence to ensure all parts of the Netherlands are present in the data.

The survey data is processed after the collection in terms of (1) removing unusable and unreliable data, (2) enriching and imputing data and, (3) checking and correcting data. In the first step, the data of a respondent is removed from the data when that particular respondent report unlikely trips or characteristics. Secondly, the data is enriched with data about the incomes of the respondents. In the
third step the data is checked if the data is complete, plausible and consistent. If this is not the case the data may be added or altered.

Finally the data is weighted to assure the representativeness. Although the survey population is supposed to be representative in itself, some groups of respondents are more likely to join the survey then other groups. Three weight factors are introduced in this process: based on (1) persons, (2) households and (3) trips. In terms of personal factors characteristics such as age, gender, province, household size, urban density and household income are used for weighting. Because the survey collects trips for one day the yearly number of trips is supposed to be a factor 365 higher. However, the trip weight factor takes irregular travel behaviour such as holidays into account.

Although the data is weighted in terms of personal, household and trip characteristics the data still suffers from some underrepresentation. It is stated that business trips, professional trips (e.g. trips made by truck drivers and other professional drivers) and trips with the purpose of touring or walking around are underreported in the data. More in general it is assumed short-distance trips are underreported, because they are more easily forgotten or assumed to be unimportant to the respondent. The actual data records retrieved from this data source are discussed in section 5.3.1.

2.3 Traffic network
The traffic network is the main means for the road authority to enable mobility and therefore also to influence traffic within their sphere of influence. The bicycle network is therefore an important aspect for the bicycle monitoring and policy making. The first part of this section presents the bicycle network viewed from a transport policy perspective (i.e. the municipality of Enschede assigned priorities to bicycle paths to cater for a high-quality bicycle network to stimulate the use of the bike) and the second part provides the translation of the network into a set of links, nodes and centroids as a model of the actual network.

2.3.1 Bike network Enschede
The bike has a prominent place in the transport network in Enschede. The traffic network for cyclists is currently improved to enhance the accessibility for intra- and interurban bicycle trips. The municipality of Enschede as road authority denoted three main focus points in the bicycle network: (1) bundled routes, (2) unbundled routes and, (3) cycling highway. This categorization was installed to ensure a high-quality bicycle network in terms of directness, connectivity, attractiveness, safety and comfort. The figure below shows the bicycle network on a map.
The red lines represent the main urban arterials for cars that also have adjacent bicycle paths and bundle several modalities on a link. The green lines represent main bicycle routes through neighbourhoods on minor roads and are assigned as main bicycle links to ensure safe and comfortable connection between the neighbourhoods and city centre (unbundled routes). There are some minor roads that have the cyclist as primary user (i.e. fietsstraat). The blue lines are the bicycle highways to be constructed. The bicycle highway is in development to carry long-distance bicycle trips between the cities in the region of Twente. On the other, minor residential roads the bike shares the road with other modalities. Ideally the bicycle flows on the network as assigned by the municipality is well monitored to have a proper view on the bicycle flows in the city to support effective policy making and evaluation. However, the municipality lacks a bicycle monitoring strategy to provide the required information from the bicycle network.

2.3.2 Regional Traffic Model
In this project a traffic network is required to enable the modelling and that can serve as information carrier for the development of the cycling database and application. The Region of Twente uses a traffic model (i.e. Regional Traffic Model (RVM)) for modelling and evaluating traffic measures on a regional scale. The use of the model is geared to cars. Although the cycling network is present in the model, the lack of bicycle data in the model prevents the use for modelling measures related to cycling. The model consists of centroids representing the origin and destination of trips and links and nodes connecting the centroids and representing the roads in a traffic network.
The model contains a fine-grained network of roads in the region of Twente. Outside this area only the major roads are modelled (the further out of the region of Twente, the coarser the network). The primary bicycle network as denoted by the municipality is largely present in the model. Some sections of the bicycle highway needed to be added to the network in the model.

2.4 Other data sources

There are various other data sources available in Enschede (e.g. data from parking garages, visual counts of car and bicycle traffic) that was not used in this project. One of the most promising upcoming data sources for travel behaviour research is the smartphone. The latest developments include a dedicated app on a smartphone that enables automatic trip registration. This data source was not included, but is mentioned here because of the potential future contribution to the field of travel monitoring.

The municipality of Enschede participated in a European project called SUNSET\(^1\) with the aim of studying the potential of changing individual travel behaviour by tracking the mobility of participants via a dedicated app on their smartphones and provide them with incentives to travel more sustainably. Resulting from the project a database with trips of participants (similar to trip characteristics of the NTS data) was collected. The data entails locations traces of these trips collected through the GPS, Wi-Fi and GSM of the smartphone, reflecting the origins, destinations, modes of transport and routes. This approach of travel data collection eliminates the human factor in trip registration compared with the NTS. Although the trip data is generated by a limited number of participants that aren’t representative for the population of Enschede (i.e. only smartphone owners could participate, self-selection as travellers interested in mobility issues joined), the data does provide route choice information (revealed preference). This can shed a light on the route choice behaviour of cyclists and the utility of the different route alternatives. The following figure provides an example of the location traces for several weeks of one specific participant.

---

\(^1\) [project website SUNSET (7 August, 2015)]
The figure shows the locations of the specific participant aggregated over several weeks. The roads the participant used and the places where the participant resides now appear. When filtering the trips by bike with a specific origin and destination (e.g. commuting) the preferred bicycle route is deduced suggesting the highest utility amongst the possible alternatives.

Although the data collected with smartphones is potentially more reliable than the reported data from the NTS, the data source is not a well-established data set yet. The technology is still in development. In a follow-up project called SMART the municipality of Enschede brings the concept to a next level implementing a system with a dedicated smartphone app allowing the general public to take part in monitoring travel behaviour and receiving rewards for sustainable travel behaviour change. In the Dutch Mobile Mobility Panel similar technology is used to retrieve trip data over a longer period of time (Thomas et al., 2014). The automated trip registration is followed by a check by the participants to complete their trip registration. In this way erroneous registrations can be corrected and missed trips can be added. Although the data resulting from this project is improving in quality (the underregistration of trips is reduced), the reliability isn’t evaluated as extensively as the NTS data yet. The data is therefore not used in this project, but can offer new insights in future applications.
3 Bicycle volumes

In this chapter, bicycle data for the city of Enschede are constructed based on the data collected at signalized intersections (intersections with traffic lights) as described in the previous chapter and is based on Veenstra et al. (2013). The municipality of Enschede collects bicycle volume data using visual counts at a number of locations and for two hours each year. This method of data collection has strong limitations for transport planning. Firstly, visual counts are a relatively expensive method of data collection. As a result, data are collected only on a limited number of locations on the network. This does not allow for constructing system wide bicycle volume data. Secondly, the visual traffic counts are conducted for a limited time period (two hours). The particular conditions under which the visual counts were executed (e.g. weather conditions, public events) may have a stronger influence on the observed numbers of cyclists than the policy measures that are evaluated (Nordback and Janson, 2010, Kidarsa et al., 2006, Dharmaraju et al., 2001). To have a better understanding of bicycle flows, a more continuous data source extended over a longer period of time is needed (Thomas et al., 2012). This way, the dynamics of cycling can be investigated and the effects of policy measures can be disentangled from external influences. In this chapter, we examine if inductive loops on bicycle lanes at signalized intersections can offer an additional data source. Inductive loops on cycling lanes are present at many signalized intersections in the Netherlands. However, the quality of this data source needs to be examined, because it is not clear a priori whether all cyclists will be detected by inductive loops, especially at high volumes (Griswold et al., 2011).

In the project the method is used to convert the traffic light data into estimates of bicycle volumes at intersections. This data now enables the construction of bicycle volume profiles at the count locations. An extensive data source for bicycle volume information can now be obtained.

The rest of the chapter is structured as follows. In Section 1, we present the visual and inductive loop data. Section 2 discusses measurement biases using inductive loop counts. In section 3, the relation between these two counts are presented, and based on the results from section 2, a generic calibration scheme for inductive loop counts is introduced. In section 4, the chapter concludes with a discussion and future work.

3.1 Data sources

The data consists of counts from detection loops, processed by the traffic light control system, and visual counts. These were gathered at several signalized intersections in the town of Enschede.

3.1.1 Inductive Loops

In Enschede approximately 50 signalized intersections are equipped with a control system that collects data from all inductive loops and signal groups. Apart from the operation of the traffic signal control, the data of the inductive loops can be used to estimate volumes at signalized intersections (U.S. Department of Transportation and Bureau of Transportation Statistics, 2000, Nordback and Janson, 2010). However, in contrast to cars, it is sometimes more difficult to detect all cyclists separately, as cyclists can cycle together or cycle in a ranging formation. In these cases, some cyclists may not be detected, because separate detections are only possible with a sufficient headway between two cyclists. Hence, when two cyclists follow each other in close range, these two cyclists will be recorded as one. As was shown in Figure 2-1 signal groups for cyclists have two inductive loops to detect incoming cyclists. Cyclists pass a distant loop first and stop just before the stop line. Near the stop line there is a second loop. Normally a cyclist occupies the stop line loop while waiting
for the traffic light. The distant loop is located approximately 15 meters prior to the stop line and detects incoming cyclists. The count data from the distant loops are used to estimate bicycle volumes. The data from the stop line loops are less suitable to estimate bicycle volumes, because individual cyclists cannot be distinguished when simultaneously occupying the stop line loop while waiting for a red light. At distant loops this problem only occurs when the queue reaches the distant loop. The latter is unlikely as queues of cyclists are generally short.

A typical inductive loop detection pattern is shown in Figure 3-1. The signal group is named ‘28’ and the accompanying stop line loop and distant loop are named ‘281’ and ‘282’ respectively. The figure illustrates that individual cyclists can be better distinguished at the distant loop than at the stop line loop. At the stop line, the loop is continuously occupied (indicated by the blue line in the center panel) when the traffic light is on red. During green (indicated by the green line in the upper panel), the loop is only occupied when a cyclist is passing. For the distant loop (lower panel), such a pattern can also be observed during red, and according to the figure, 9 cyclists passed in this time period. However, the figure shows that even for distant loops continuous occupation times are sometimes larger (in this case the 6th and 7th detection during the red phase) than may be expected from the speed of cyclists. These particular detections could be due to slow cyclists occupying the loop for a longer period of time or due to a pair or cluster of cyclists being detected as one.

In this case we used 15-minute aggregates of the inductive loop counts. Future improvements in counting bicycles may arise from investigating the total loop occupancy in the 15-minute interval and its relation to the actual number of passing cyclists. This may be another indicator for the expected number of passing cyclists. Instead of using 15-minute aggregates the actual detection patterns could be investigated. Using for example the occupancy time of single detections at the distant loop and the uninterrupted occupancy time of the stop line loop during green (i.e. the time it takes for the queue to resolve after the traffic light turns to green) may provide additional information to make a more accurate estimation of the number of passing cyclists. However, this requires a more thorough data collection process of simultaneously collecting detection patterns and for example video images, to enable a comparison of the length of single detections and the actual number of bicyclists during that specific detection. Moreover, it is still unclear if the accuracy of the counts will actually increase using detection patterns instead of 15-minute aggregates.

### 3.1.2 Visual Counts

Visual counts were executed at three signalized intersections, as shown in Figure 3-2. These intersections serve moderate to high bicycle flows. These were selected based on the assumption that accurate estimates of high volumes using inductive loops is especially challenging. Signalized Intersection (SI) 2 and 6 have on average the highest peak hour volumes, because these serve bicycle flows between the town center and the main employment and educational areas in the western part of Enschede.
In total, 13 hours of visual counts were carried out during both peak and off-peak periods on several weekdays in the months March, April and May 2012. In this way, possible differences in cycling behavior throughout the day were incorporated. Visual counting was performed in 15 minute intervals. This interval length is typical for volume studies: it is sufficiently long to reduce random variation in demand, while it still can be used to study short term dynamics in bicycle volumes.

In Table 3-1, a comparison is made between bicycle volumes at important intersections during a typical morning rush hour peak. According to the table, the median volume lies around 50 cyclists per 15 minutes, which can be considered as a typical volume for the main corridors in Enschede. For some intersections, peak volumes are between 50 and 100 counts per 15 minutes. Only for the two locations that serve cyclists coming from the city center and going to the western part of Enschede, volumes exceed 100 counts per 15 minutes.

Table 3-1: Comparison of morning peak (March 28, 2012) inductive loop counts of bicycles at major intersections in Enschede

<table>
<thead>
<tr>
<th>Time</th>
<th>SI 1</th>
<th>SI 2*</th>
<th>SI 3</th>
<th>SI 4</th>
<th>SI 5</th>
<th>SI 6</th>
<th>SI 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-8:15</td>
<td>69</td>
<td>87 (120)</td>
<td>53</td>
<td>46</td>
<td>63</td>
<td>82</td>
<td>23</td>
</tr>
<tr>
<td>8:15-8:30</td>
<td>86</td>
<td>102 (170)</td>
<td>34</td>
<td>31</td>
<td>72</td>
<td>87</td>
<td>17</td>
</tr>
<tr>
<td>8:30-8:45</td>
<td>60</td>
<td>118 (209)</td>
<td>25</td>
<td>39</td>
<td>55</td>
<td>102</td>
<td>11</td>
</tr>
<tr>
<td>8:45-9:00</td>
<td>52</td>
<td>118 (187)</td>
<td>18</td>
<td>32</td>
<td>49</td>
<td>98</td>
<td>12</td>
</tr>
<tr>
<td>9:00-9:15</td>
<td>24</td>
<td>80 (105)</td>
<td>17</td>
<td>24</td>
<td>33</td>
<td>66</td>
<td>12</td>
</tr>
<tr>
<td>9:15-9:30</td>
<td>28</td>
<td>63 (77)</td>
<td>13</td>
<td>14</td>
<td>31</td>
<td>64</td>
<td>14</td>
</tr>
</tbody>
</table>

* between brackets: visual counts at the intersection for the specific morning peak (March 28, 2012)

3.2 Measurement biases in using inductive loop detectors
The visual counts can be used to calibrate counts from inductive loop detectors. In the next section, an empirical relation between inductive loop detections and the actual bicycle volumes is derived.
through the visual counts for the three selected intersections. In theory, this empirical relation can serve as a calibrator for inductive loops counts from other signalized intersections and time periods. However, the amount of data used in the empirical relation is limited. To determine the validity of the empirical relation for other cases, an estimation of expected deviations between visual and inductive loop counts is given a priori in this section.

Two main types of measurement biases using inductive loop detectors are expected. The first type is most likely proportional to bicycle volume. This kind of bias can have various causes. Some of these may lead to an underestimation of inductive loop counts. Individuals that cycle together, for example, arrive at the inductive loop at the same time and are therefore counted as one. Another example is cyclists that miss the loop entirely, because they cycle on the sidewalk or on the main road. Other causes may lead to an overestimation of inductive loop counts, such as cars or cyclists from the opposite direction crossing the detection loop. It is not clear to what extent these deviations are related with volume, but in general it can be assumed that these deviations are rather small. This bias is assumed to be proportional to the bicycle volume and can therefore be described as a percentage of the total volume. The causes can have opposite effects, but the net effect is probably an underestimation by inductive loops. However, this underestimation is probably not more than a few per cent.

The second type of bias arises when the headway between cyclists is very small. When two cyclists follow each other within close range, inductive loop detectors will detected them as one. If we assume a random arrival process, or any other arrival process for that matter, the headway between two successive arrivals will in some cases be too small for the inductive loop to distinguish between two cyclists. The minimum time headway \( (h_{\text{min}}) \) beyond which two cyclists can be detected separately is:

\[
h_{\text{min}} = \frac{(l_{\text{bike}} + l_{\text{loop}})}{v_{\text{bike}}}
\]

with \( l_{\text{bike}} \) and \( l_{\text{loop}} \) the length of the bike and inductive loop respectively, and \( v_{\text{bike}} \) the velocity of a cyclist. Typical values are 2 meters for \( l_{\text{bike}} \), 1 meter for \( l_{\text{loop}} \) and 5 m/s (18 km/h) for \( v_{\text{bike}} \), resulting in a minimum time headway of 0.6 seconds. The probability that the next cyclist is detected separately from the previous cyclist \( (P_{\text{det}}) \) is thus:

\[
P_{\text{det}} = P(h > h_{\text{min}})
\]

The expected number of detections for \( N \) cyclists \( (N_{\text{det}}) \) is then:

\[
N_{\text{det}} = 1 + \sum_{N}^{2} P(h > h_{\text{min}}) \approx N * P(h > h_{\text{min}})
\]
\[ s(N_{det}) = \sqrt{(N_{det} \cdot (1 - P(h > h_{\text{min}})))} \]  

Equation 3-4

When we assume a random arrival process, the time headways have a negative exponential distribution. According to the corresponding cumulative distribution:

\[ P(h > h_{\text{min}}) = \exp\left(\frac{-h_{\text{min}}}{h_{\text{avg}}}\right) \]  

Equation 3-5

For 15 minute cycle volumes \((N_{15})\), the average time headway \((h_{\text{avg}})\) in seconds is given as:

\[ h_{\text{avg}} = \frac{15 \times 60}{N_{15}} \]  

Equation 3-6

For small volumes, the average time headway will be relatively large, and the corresponding probability \(P(h > h_{\text{min}})\) close to 1. Only for (very) large volumes, the rate of underestimation will not be negligibly small as \(h_{\text{avg}}\) decreases. However, Equation 3-5 gives an upper limit. The number of bicycles that are not detected will be larger in reality, because cyclists do not necessarily arrive randomly. Especially when traffic lights are in close proximity of each other, clusters of cyclists will travel from the previous to the next traffic light. Within these limited arrival time windows, the distribution of time headways can still have a random character, because cyclists do not travel with the same speed. However, the average headway in this distribution will be much smaller. Therefore, Equation 3-5 can be adjusted to:

\[ P(h > h_{\text{min}}) = \exp\left(\frac{-h_{\text{min}}}{c \cdot h_{\text{avg}}}\right) \]  

Equation 3-7

with \(0 < c \leq 1\), indicating the fraction of time (clusters of) cyclists pass the inductive loop. For isolated traffic lights, \(c\) is expected to be close to 1. For a (dominant) flow of cyclists traveling between traffic lights in close proximity, \(c\) is expected to be close to the green fraction of the previous traffic light.

### 3.3 Comparison of visual and inductive loop counts

In Figure 3-3, the visual and inductive loop counts are plotted for the three selected intersections. The dashed line shows the theoretical cumulative distribution according to the random arrival process (Equation 3-5) and therefore the rate of underestimation by the inductive loop due to a fraction of small headways in a random arrival process. As expected, this theoretical limit does not sufficiently explain the rate of underestimation. The best fit between visual and inductive loop counts is found for the time fraction \(c = 0.24\), indicated by the solid line. The dotted lines give the uncertainty band for this fit, i.e. the standard deviation (Equation 3-4) with respect to the expected value.
Figure 3-3 shows that the correlation between visual and inductive loop counts is generally quite good ($R^2 = 0.96$). The root-mean-square (RMS) in the residuals (differences between observations and values according to the best fit) is on average 8% of the volume. This is only slightly larger compared with the theoretical standard deviation of about 6%, which is quite constant over the volume range. According to the figure, other effects, as described in Section 3, are limited in magnitude.

The time fraction of random arrivals, $c$, is quite small for the selected intersections. However, this fit is mainly based on the intersection with the highest observed volumes (i.e. SI 2). This intersection serves a large flow of students and university employees in the morning rush hour. This flow mainly consists of cyclists arriving from the center, which typically travel in clusters, because the previous traffic light is situated only 400 meters (0.25 mi) upstream (i.e. SI 6 in Figure 3). For other more isolated signalized intersections, smaller rates of underestimation and therefore higher values of $c$ are expected. However, this is still to be examined.

For a generic calibration tool, deviations in $c$ are only relevant for large volumes. For volumes around and below 50 cyclists in 15 minutes (200 cyclists per hour or less), the inductive loop counts correspond well with visual counts, and structural differences are typically below 10%. According to Table 1, flows at important intersections have typically around 50 inductive loop detections in 15 minutes, except from a few peak moments and / or intersections. For inductive loop counts between 50 and 100 per 15 minutes our model can convert the counts into actual bicycle volumes. In middle-sized cities in the Netherlands, with relatively low bicycle volumes (up to about 500 bicycle per hour), inductive loop counts in combination with our model will represent actual counts quite well. In these cases, structural deviations are often negligibly small. However, in major cities in the Netherlands, actual bicycle flows at major bicycle lanes during peak hours are much higher. In Amsterdam, for
example, at the 10 busiest bicycle lanes 1500 to 2000 bicycles pass per hour during evening rush hours, and a large part of the road network in Amsterdam has volumes over 500 cyclists per hour during rush hours (Nordback and Janson, 2010, Kidarsa et al., 2006, Dharmaraju et al., 2001). For the large and/or most popular bicycle cities in the Netherlands, our model may no longer be suitable to calibrate inductive loop counts and may thus not offer a useful data collection method.

3.4 Discussion

It was shown that inductive loop data at signalized intersections can be used to accurately estimate bicycle volumes. In the town of Enschede, visual counts of cyclists were made at three signalized intersections equipped with inductive loops on the cycle path. Visual counts could thus be used to calibrate counts from inductive loops. Around and below 50 counts per 15 minutes (200 cyclists per hour), inductive loop counts correspond well with visual counts. For higher volumes, the inductive loop counts underestimate actual bicycle volumes, which can be explained by assuming a random arrival process of cyclists within a certain timespan.

The calibration coefficient, i.e. the timespan in which cyclists arrive, may not be the same for all intersections. Depending on the proximity of other signalized intersections, cyclists arrive in clusters, especially in the urban environment. Therefore, the smaller the proximity of two signalized intersections, the shorter the timespan in which cyclists arrive will be. Further research is needed to study the effect of proximity of intersections in more detail. However, for the town of Enschede, it can be concluded that inductive loop counts can already be used in most cases, because most intersections have (peak) volumes around or below 50 cyclists per 15 minutes. For inductive loop counts between 50 and 100 per 15 minutes the inductive loop counts need to be calibrated to obtain the actual bicycle volumes.

The data from inductive loops at signalized intersection can be used to estimate bicycle volumes and can provide an extensive and valuable source of data concerning bicycle volumes (DIVV, 2010). As it is a continuous, widely available and low-cost data source, the dynamics in bicycle volumes (e.g. due to weather conditions and public events) can be investigated. This data source and the understanding of the dynamics in bicycle volumes will be of great value to practitioners, road authorities and local administrations in encouraging cycling as a mode of urban transport and improving the accessibility and sustainability of the urban transport system.
4 Bicycle flow estimation

The importance of cycling in the urban areas is increasing, as sustainable modes of transport are the focus of urban transport policy nowadays. However, municipalities, who are responsible for bicycle planning, generally lack procedures that include system wide bicycle volume data in contrast to procedures for collecting, summarizing, and disseminating motor vehicle traffic volumes. Municipalities typically use aggregate data from the National Travel Survey, combined with visual counts. The lack of data on bicycle volumes and flows hampers municipalities to plan and improve bicycle facilities. As a result, local transport models typically are not very well-suited to model bicycle trips.

On the other hand, the current traffic management systems deployed in the urban environment collect more and more data. As was presented in chapter 3 a signalized intersection provides a continuous source of traffic volume data to provide information about the dynamics of bicycle volumes at various locations in the network. However the traffic light data lacks the trip characteristics of the passing cyclists. The NTS data provides information about the characteristics of bicycle trips (e.g. origin-destination pairs, timing, trip purpose, trip distance) to construct an origin-destination matrix for the Twente Region, but the data source fails the critical mass to construct sensible bicycle flows (on a yearly basis about 800 trips are registered in Enschede). Combining the two sources may provide the best of both worlds and may give an insight in urban bicycle flows for the municipality of Enschede.

The objective of this part of the study is to combine bicycle volume data from traffic lights and origin-destination pairs of bike trips from the NTS to estimate urban cycling flows in the city of Enschede and to evaluate sensitivity of these bicycle flows to variations in the input (i.e. zonal production and attraction estimates, initial OD-matrix, OD-matrix estimation procedure and traffic assignment) to find a method for constructing bicycle flows in the traffic network of Enschede.

This chapter first describes the data sources (i.e. bicycle volumes and travel behaviour data) available for this study and an initial analysis is conducted to retrieve aggregated measures of volumes and trip rates in Enschede. Secondly the transport network is selected as the carrier of the urban cycling flows. The third section the procedure of estimating cycling flows is described. The procedure consists of (1) constructing initial OD-matrices, (2) assigning the matrices to the network and (3) calibrating the flows with the bicycle counts. In the fourth section the results are presented. Based on the trip generation, the trip length distribution and the comparison with the bicycle counts the accuracy and reliability of bicycle flows resulting from the various initial OD-matrices is discussed.

4.1 Data sources

In the research we used and combined two different data sources (1) bicycle detections at signalised intersections in the city of Enschede and (2) bicycle trips in Enschede extracted from the National Travel Survey (NTS). We projected these on a fine-grained traffic network from the transport model of the region of Twente.

4.1.1 Bicycle detections

The data from traffic lights in Enschede consist of inductive loop detections at bicycle paths leading up to signalized intersections. Only the count locations with a sufficient amount of reliable data were selected to be used in this analysis. Count locations with gaps in the data were excluded.
Figure 4-1: count locations in Enschede used in the bicycle flow calibration

At these intersections the actual detections per 15-minute interval were processed into estimations of bicycle volumes according to the method described in chapter 3.

The processed data was then categorized according to the day of the week and the time of day. Before constructing average volumes for the time periods events such as unreliable data and days with road works and specific scheduled events were filtered out. The data wasn’t corrected for any weather events. Average volumes were calculated for the following time periods:

- Workday (sum of the average number of bicycles passing during the day for all workdays combined)
- Weekend day (sum of the average number of bicycles passing during the Saturday)
- Morning peak (sum of the average volumes on workdays from 6AM to 10AM)
- Evening peak (sum of the average volumes on workdays from 3PM to 7PM)

For all signalized intersections with bicycle counts the following figures were constructed. The figures depict the volume profiles for the various time periods of one of the main intersections in the Enschede urban area for cyclists, situated between the city center and the main commercial and industrial area also containing the university.
Figure 4.2: example of bicycle profiles (intersection Hengelosestraat and Singel)

The figures show the traffic light data enables the presentation of the dynamics in bicycle volumes.

4.1.2 Bicycle trips from the National Travel Survey

The NTS data does provide information about the origins and destinations, timing and purpose of the trips and the trip maker. The NTS editions from 2004 to 2013 were used to increase the critical mass of trips in this dataset. On a yearly basis approximately 500 cycling trips are registered in Enschede. Combining the consecutive editions resulted in a set of 8216 bicycle trips. The editions from 2003 and before were ignored because of a change in the survey set-up. The selected editions were combined and the bicycle trips originating in and/or going to Enschede were selected to study the aggregate characteristics of bicycle trips in Enschede.

Average trip rate and trip length distribution of citizens of Enschede

From the National Travel Survey aggregate statistics about cycling can be deduced and projected on the inhabitants of Enschede. For example, the number of bike trips for various age categories and trip purposes. The following figure shows the demography is an important factor in the bike trip production (based on the entire Dutch population). On average a person makes 0.9 bike trips per day. However this rate varies between the age groups. For children the bike is the main source of transportation. From the age of 25 to approximately 65 the daily bike trip rate remains rather
constant at 0.75 bike trips per day. After 65 the trip rate drops. This pattern is primarily caused by bike trips to school, indicating that schools are main attractors for bike trips. When zooming in on Enschede, a similar picture arises. However the average bike trip rate is slightly higher with 1.03. This figure however appears to be less smooth due to low number of cases more random variations in trip rates occur.

![Figure 4-3: daily trip rate by bicycle in The Netherlands (left) and Enschede (right) extracted from the NTS](image)

The average trip rate can be used to estimate the number bicycle trips produced by the zones based on the number of inhabitants of the zones. To account for the higher trip rate of children the age distribution within the zones should be used to estimate the trip production. However according to demographic data from the Statistics Netherlands the distribution of age is comparable for all neighborhoods in Enschede. This implies all neighborhoods will ‘produce’ about 1 bicycle trip per inhabitant.

Evidently, the trip attraction of zones is related to the production. However, one needs the exact locations or at least the zone where the attractor is located. To make an estimate of the trip attraction one should know whether or not there is a school in that specific zone and the size of the school, but also if there is a supermarket or other retail areas, jobs, recreational areas and other services in the specific zone. In this case the number of jobs of a zone is stored in the RVM. Data about the other attractors is lacking. Moreover the NTS can also provide a trip length distribution based on the reported trip lengths of bike trips. This measure can then be used to calibrate OD matrices for bicycles for Enschede. From the figure one can conclude that for this research that commuting trips by bike are the most important bike trips. When aligning with the bike counts at intersections there is a higher change long-distance trips pass by these count locations. Therefore the long-distance bike trips are the most important. Moreover, bike trips to a primary school or a supermarket are more often short distance trips not passing by the count locations. Although they are important to get a proper overview of bicycle flows in the urban environment, they are less ‘visible’ in the transport system. We therefore assume that having incorporated the number of inhabitants and the number of jobs at a centroid in the network accounts for most of the longer-distance bike trips that can be detected at the count locations.
Figure 4-4: trip length distribution of bicycle trips in Enschede

Trip production and attraction per postal zone
From the NTS the total production and attraction for the various postal zones in Enschede were extracted. This trip generation is compared with the socio-economic contents of the postal zones as extracted from the RVM and presented in the following table.

Table 4-1: comparison of trip generation based on the NTS and the RVM

<table>
<thead>
<tr>
<th>PC4</th>
<th>NTS</th>
<th>RVM</th>
<th>Estimated bike trips (weight factor)</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>production</td>
<td>attraction</td>
<td># centroids</td>
<td># inhabitants</td>
</tr>
<tr>
<td>7511</td>
<td>498</td>
<td>507</td>
<td>25</td>
<td>8055</td>
</tr>
<tr>
<td>7512</td>
<td>81</td>
<td>83</td>
<td>13</td>
<td>7219</td>
</tr>
<tr>
<td>7513</td>
<td>108</td>
<td>112</td>
<td>13</td>
<td>4851</td>
</tr>
<tr>
<td>7514</td>
<td>115</td>
<td>113</td>
<td>12</td>
<td>4611</td>
</tr>
<tr>
<td>7521</td>
<td>198</td>
<td>202</td>
<td>10</td>
<td>8525</td>
</tr>
<tr>
<td>7522</td>
<td>186</td>
<td>184</td>
<td>18</td>
<td>6716</td>
</tr>
<tr>
<td>7523</td>
<td>167</td>
<td>167</td>
<td>17</td>
<td>12736</td>
</tr>
<tr>
<td>7524</td>
<td>17</td>
<td>18</td>
<td>11</td>
<td>3558</td>
</tr>
<tr>
<td>7525</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>302</td>
</tr>
<tr>
<td>7531</td>
<td>152</td>
<td>153</td>
<td>13</td>
<td>8256</td>
</tr>
<tr>
<td>7532</td>
<td>63</td>
<td>60</td>
<td>18</td>
<td>8471</td>
</tr>
<tr>
<td>7533</td>
<td>61</td>
<td>58</td>
<td>5</td>
<td>4822</td>
</tr>
<tr>
<td>7534</td>
<td>191</td>
<td>186</td>
<td>15</td>
<td>13357</td>
</tr>
<tr>
<td>7535</td>
<td>104</td>
<td>100</td>
<td>5</td>
<td>3177</td>
</tr>
<tr>
<td>7536</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>254</td>
</tr>
<tr>
<td>7541</td>
<td>100</td>
<td>97</td>
<td>10</td>
<td>3627</td>
</tr>
<tr>
<td>7542</td>
<td>146</td>
<td>146</td>
<td>9</td>
<td>9971</td>
</tr>
<tr>
<td>7543</td>
<td>107</td>
<td>107</td>
<td>10</td>
<td>7410</td>
</tr>
<tr>
<td>7544</td>
<td>191</td>
<td>190</td>
<td>17</td>
<td>15636</td>
</tr>
<tr>
<td>7545</td>
<td>255</td>
<td>259</td>
<td>14</td>
<td>14808</td>
</tr>
<tr>
<td>7546</td>
<td>144</td>
<td>138</td>
<td>8</td>
<td>8381</td>
</tr>
<tr>
<td>7547</td>
<td>40</td>
<td>37</td>
<td>11</td>
<td>461</td>
</tr>
<tr>
<td>7548</td>
<td>65</td>
<td>61</td>
<td>6</td>
<td>3404</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Estimated bike trips (weight factor)</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>2999</td>
<td>2988</td>
<td>62119,5</td>
</tr>
<tr>
<td>262</td>
<td>158608</td>
<td>0,0168</td>
</tr>
<tr>
<td>79685</td>
<td>158798</td>
<td>0,019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Estimated bike trips (weight factor)</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>31040,5</td>
<td></td>
<td>0,0260</td>
</tr>
<tr>
<td>12263,25</td>
<td></td>
<td>0,041</td>
</tr>
<tr>
<td>7277,75</td>
<td></td>
<td>0,011</td>
</tr>
<tr>
<td>4942,75</td>
<td></td>
<td>0,022</td>
</tr>
<tr>
<td>6606,75</td>
<td></td>
<td>0,017</td>
</tr>
<tr>
<td>9972,25</td>
<td></td>
<td>0,015</td>
</tr>
<tr>
<td>3332</td>
<td></td>
<td>0,005</td>
</tr>
<tr>
<td>256</td>
<td></td>
<td>0,035</td>
</tr>
<tr>
<td>6976</td>
<td></td>
<td>0,022</td>
</tr>
<tr>
<td>8194,25</td>
<td></td>
<td>0,008</td>
</tr>
<tr>
<td>3890,5</td>
<td></td>
<td>0,016</td>
</tr>
<tr>
<td>10798,75</td>
<td></td>
<td>0,018</td>
</tr>
<tr>
<td>2661,25</td>
<td></td>
<td>0,039</td>
</tr>
<tr>
<td>221,5</td>
<td></td>
<td>0,005</td>
</tr>
<tr>
<td>4221,75</td>
<td></td>
<td>0,024</td>
</tr>
<tr>
<td>7794,25</td>
<td></td>
<td>0,019</td>
</tr>
<tr>
<td>7757</td>
<td></td>
<td>0,014</td>
</tr>
<tr>
<td>13000,5</td>
<td></td>
<td>0,015</td>
</tr>
<tr>
<td>12517,5</td>
<td></td>
<td>0,02</td>
</tr>
<tr>
<td>7091,75</td>
<td></td>
<td>0,02</td>
</tr>
<tr>
<td>6311,75</td>
<td></td>
<td>0,006</td>
</tr>
<tr>
<td>3425</td>
<td></td>
<td>0,019</td>
</tr>
</tbody>
</table>
The table above presents the socio-economic content of all postal zones (PC4) in Enschede. The second and third column present the production and attraction of the postal codes as retrieved from the NTS. The table shows the production and the attraction are comparable. From this we assume the full day OD-matrix can be represented as a symmetrical matrix (NB this is not necessarily true when considering a morning or afternoon OD-matrix). The number of inhabitants and jobs was retrieved from the RVM. From these zonal characteristics a coarse estimation of the daily number of bike trips produced and attracted by the postal zone was calculated. Although more characteristics should be incorporated (e.g. school places and shopping area) to properly estimate bicycle trip generation, we assumed only incorporating the number of inhabitants and jobs is sufficient for this purpose, because the weight factor is mainly used to distribute the trips of a postal zone among its centroids. The trip generation was computed by multiplying the number of inhabitants by 0.75 (bike trip production) and multiplying the number of jobs by 0.5 (trip attraction). This estimation was used as a weight and scaling factor in the generation of the initial OD-matrices. Moreover, comparing the trip production and attraction based on the NTS and weight factor from the RVM can provide more insight in zones with a particulary high or low trip rate. This could be a sign of a structural under- or overestimation of the weight factor and estimated bike trips in the RVM. Comparing the trip production based on NTS and the estimation based on the contents of the centroids the difference is a factor 50. The OD-matrix extracted from the NTS should therefore be multiplied by factor 50 to approach the actual daily traffic flows in the network.

There are some postal zones where the weight factor is not in correspondence with the production and attraction as retrieved from the NTS. For example, the weight factor for postal code 7511 underestimates the bike trip generation because this postal zone contains the main shopping area and therefore attracts more shopping trips than estimated based on inhabitants and jobs. Another notable difference occurs in the main commercial area (postal zone 7547). This area is located on the edge of Enschede. Because of its location and its accessibility by car the bike has a small share in the modal split of trips from and to this area. This implies that more variables than just inhabitants and jobs should be included in the weight factor. The location, type of jobs and the amount of shopping area could be included. Moreover, when investigating Figure 4-3 a large share of bike trips has the purpose of going to school. Therefore also the number of school places in a zone should be used as an indicator of trip generation. However, this information is not available within the scope of this research.

4.2 Traffic network selection

We extracted the network from the regional traffic model from the Region of Twente. This model is intended to be used to execute modelling studies in policy making and evaluation. The network in the model contains both the car, public transport and the slow traffic modes (bike and foot). In this research we focused on using the bicycle network. The advantage of using the RVM network over other networks is reliability and connectivity. Within the network it is assumed the RVM network is better than for example the OpenStreetMap network. Moreover the RVM network contains road characteristics and was specifically designed for the region of Twente.

The RVM network compares to the actual traffic network in the sense that all major arterials and collector roads are present in the model. Minor rural roads and roads in residential areas (living streets) were not included in the RVM network. The socio-economic information of areas is stored in the centroids. A centroid in the RVM network represents a rather homogenous composition of an
area in terms of inhabitants and jobs. Depending on the urban density there can be many centroids within a single postal zone.

Figure 4-5: coverage of RVM links compared with entire network (left) and density of centroids compared with postal zones (right)

To the RVM network we added the counts at the signalized intersections. In accordance with the availability and reliability of the data of the signalized intersections we used in total 37 inductive loop detectors throughout the traffic network of the city of Enschede.

4.3 Method
The process of constructing bicycle flows on the network of Enschede follows several steps: (1) constructing an initial od-matrix, (2) assigning the od-matrix to the network and (3) calibrating the resulting bicycle flows to the average bicycle counts in the network.

4.3.1 Constructing an initial OD-matrix
The first step of generating bicycle flows is constructing an initial origin-destination matrix. This represents the initial relation between the production and attraction of zones in the network. The initial OD-matrix can have several levels of detail as a result of the available data. When extensive origin-destination data is available, it is evident to use this data to construct the initial matrix. However, when this data is not available or only aggregate data is present, an initial OD-matrix can still be constructed. In this section we describe three approaches ordered in increasing levels of OD-data availability: (1) a matrix generated using a gravity model (based on the weight factor for the trip generation and the trip length distribution), (2) a NTS matrix (using the OD-data from the NTS) and (3) a mixed matrix (using the trip generation form the NTS and the distribution from the gravity model).

Gravity model
A matrix was estimated based on the estimated trip ends of all RVM zones and the trip length distribution as extracted from the NTS for the Twente Region. Via trial-and-error a top-lognormal and a discrete trip distribution function was estimated with which the trip length distribution as extracted from the NTS was reproduced. The following figure shows the trip length distribution for (1) all bike trips by Enschede citizens and (2) all bike trips either originating or ending in Enschede.
Figure 4-6: trip length distribution of bike trips of citizens of Enschede and trips from and to Enschede

The figure shows some differences between the distributions. The citizens of Enschede make more short distance trips (under 7.5 kilometers) than travellers from and to Enschede. This implies Enschede attracts cyclists from surrounding municipalities. When investigating trips originating or ending in Enschede there is a higher share of bike trip of around 10km. This is caused by bike commuters working in Enschede and living in the surrounding cities (primarily Hengelo). For the two trip length distributions we estimated distribution functions that return the trip length distributions as extracted from the NTS. Firstly, via trial-and-error a top lognormal distribution function was estimated by trying various configurations of parameters ($\alpha$, $\beta$ and $\gamma$), however the trip length distribution wasn’t fully reproduced. Therefore we chose to estimate a discrete distribution function to better incorporate the specific spatial configuration of Enschede (i.e. Enschede is a regional center). The discrete distributions function was estimated by categorising network distances between origins and destinations into distance bins (with a size of 1km). For every distance bin a distribution value is estimated. The collection of distribution values of all bins now reflects the distribution. After optimizing the discrete distribution functions we could reproduce the various trip length distributions accurately.
The figure shows the relative attractiveness of the various distance bins for three distribution functions. The top lognormal distribution does not fully resemble the discrete distribution functions. The top lognormal cannot mimic the trip length distribution of either the citizens of Enschede or the trip from and to Enschede. Therefore we used the trip length distribution of trip originating from or destined for Enschede (i.e. DISCRETE OD Enschede) for the estimation of the bike distribution function to best estimate the trip length distribution of the trips in the bicycle network of Enschede. In this way the share of long distance bike trips is then overestimated in general, but the trip length distribution is expected be more more inline with the traffic counts in Enschede.

The approach of using a gravity model to estimate an initial OD-matrix shows that every trip length distribution available can be used to estimate an OD-matrix. Especially when using a discrete distribution function any given trip length distribution can be reproduced. This implies that this approach of generating an initial OD-matrix can take the specific spatial configuration and the resulting trip length distribution into account. If no specific information about trip length distributions or an actual OD-matrix is present, this approach can be used to make an estimate of the distribution. However, in this approach all centroids use the same distribution function. This is not necessarily true in all cases, because the origins and destinations are not distributed homogeniously over space. A better approach in this sense is to use the actual OD-pairs from the NTS (if available).

**NTS matrix**

The second, more straightforward approach of estimating bicycle flows in Enschede is to extract the bike trips in and around Enschede from the NTS and to use the actual OD-pairs instead of estimating them using a gravity model. The NTS editions from 2004 to 2013 were combined and from these editions all bike trips with either the origin or the destination in Enschede or one of the surrounding municipalities were selected. The selection holds 8216 trips.
In the NTS the origins and destinations are represented in terms of postal code. To be able to align the trips in the NTS with the RVM, all centroids in the RVM are assigned to one postal code (N.B. one postal code can hold multiple RVM centroids). The trips in the NTS are assigned to all RVM OD-pairs according to the corresponding postal codes as stated in the NTS and weighted and redistributed according to the estimated production and attraction of the specific RVM zone. For this procedure we used the information present in the RVM. The weight factor was composed from the number of inhabitants and the number of jobs within a centroid (according to the RVM). This weight factor is now used to redistribute trips within one specific postal zone and specifically not between postal zones. In this way the sum of the RVM OD-matrix equals the amount of trips as extracted from the NTS. The procedure to estimate the number of trips between all RVM zones is presented in the following equations:

\[ T_{ij} = T_{PC_i,PC_j} \cdot f_i \cdot f_j \]  

**Equation 4-1**

Where:
- \( T_{ij} \) = trips between RVM centroid \( i \) and \( j \)
- \( T_{PC_i,PC_j} \) = trips between postal zone containing RVM centroid \( i \) and postal zone containing RVM centroid \( j \) (as used in the NTS)
- \( f_i \) = multiplication factor for RVM centroid \( i \)
- \( O_i \) = weight factor of RVM origin \( i \)
- \( \sum_{PC_i} O_i \) = sum of all weights of RVM origins with the same postal code as \( i \)
- \( f_j \) = multiplication factor for RVM centroid \( j \)
- \( D_j \) = weight factor of RVM destination \( j \)
- \( \sum_{PC_j} D_j \) = sum of all weights of RVM destinations with the same postal code as \( j \)

The OD-matrix contains 8216 trips for Enschede and its surrounding area. The matrix now needs to be expanded to account for the entire population of the Enschede region. This was done by multiplying the OD-matrix by a factor 50. For the timespan of the extracted trips of the NTS (i.e. 2004 to 2013), Enschede had approximately 155,000 inhabitants. These inhabitants make on average 1 bike trip per day. In the NTS in total 3,000 trips originating from or going to Enschede are present. This implies the multiplication factor should be 50. The resulting OD-matrix now represents the OD-relations of an average working day for Enschede and its direct surroundings (including trips to and from Hengelo, Borne, Haaksbergen, Oldenzaal and Losser).

Using the trips from the NTS specific OD-relations are now included in the initial OD-matrix. The initial OD-matrix can be further specified using the available trip characteristics in the NTS, for example to compose a matrix for a specific trip purpose or for a specific period of the day. However, the NTS contains a limited number of trips, not allowing for further specification without losing a critical mass of trips and the accompanying loss of reliability. Therefore we only use the OD-matrix containing all trips in Enschede.
Mixed matrix
The third matrix constructed is a mixture between the gravity model matrix and the NTS matrix. In this OD-matrix the trip generation from the NTS is combined with the trip distribution from the gravity model matrix. This matrix is introduced to represent a case where the trip generation can be estimated more accurately (assuming that the trip generation is represented by the NTS more accurately than the weight factor as used in the gravity model) for example a case where more socio-economic information of neighbourhoods is available.

4.3.2 Matrix assignment
The various initial OD-matrices are assigned to the network according to an all-or-nothing assignment procedure. It is assumed that bicyclists choose the shortest route to their destination. Opposite to route choice behaviour of car drivers, bicyclists are not affected by speed limits and therefore we assume cyclists adopt their desired speed at every link in the network. This implies only distance is taken into account in route choice. We assume characteristics such as comfort are negligible in the assignment. This assumption is tested and presented in section 4.4.4. Choosing for distance as the sole driver in route choice also reduces the complexity on the assignment procedure. As long as the route choice behaviour of cyclists is unclear, we choose the most straightforward assignment procedure using only distance.

4.3.3 Matrix calibration techniques
The objective of this exercise is to estimate an OD-matrix that best resembles the bicycle flows in the traffic network. As the information from the various sources will never perfectly align, a calibration approaches should be used to find an OD-matrix that is the optimum between the information provided by the traffic light data and the relations extracted from the NTS. In this research 2 approaches are described: (1) entropy maximisation, (2) structured matrix computation

Entropy maximization
The entropy maximization approach is based on a procedure of maximizing the likelihood using the available data to find the most likely macro state, being the one that can be composed of the most combinations of micro states. For example, in the traffic domain there are many travellers making trips in a traffic network. To get an overview of the performance of the transport system, aggregate measures of variables in the system can be calculated. These aggregates (i.e. macro states) are composed of the trips of individuals (i.e. micro states) and can be expressed in terms of trip production and attraction of an area (i.e. meso state). There are many combinations of micro states or meso states that result in the same macro state. The most likely meso state is the one that can be composed by the most combinations of micro states and complies with the macro state constraints.

The following equation reflects the number of micro states associated with the meso state:

\[ W \left\{ T_0 \right\} = \frac{T!}{\prod_i T_i !} \]

Equation 4-2

To calculate the most likely meso state, one needs to identify the \( T_{ij} \) that maximizes \( W \). When now count values are introduced some additional constraints can be formulated. The problem is subject to:
\[ V_a - \sum_{y} T_{ij} p_{ij}^a = 0 \]

Equation 4-3

Where:
- \( V_a \) = count value \( V \) on link \( a \)
- \( p_{ij} \) = proportion of \( T_{ij} \) that uses link \( a \)

In this case the entropy maximization method may not be the most efficient method to use. As the method investigates the likelihood of possible matrices this specific case of a 4478 x 4478 matrix with 37 count locations generates an enormous number of potential matrices to investigate. A method not requiring matrix identification is more practical in this case.

**Structured matrix computation**

Structured matrix computation is a method to gradually bring the OD-matrix in line with the information from the traffic counts. The OD-matrix is updated according to the traffic counts in the network to make the loads reflect the counts. In this method the number of trips of OD-pairs passing by a count location is adjusted in accordance with the difference between the traffic loads (resulting from an AON assignment of the OD-matrix) and the counts. A weight factor can be used to reflect the reliability of a count value when adjusting the number of trips of an OD-pair. This method was used in this case as it is more practical to apply with the low number of counts relative to the number of links in the network. The procedure is described in more detail in the next section.

4.3.4 Matrix calibration procedure

**Counts at signalized intersections**

Constructing the average count values for the various count locations in the network (see Figure 4-1) was done through a number of filtering and sorting procedures. First of all, the intersections in Enschede equipped with distant loop detectors at the bicycle path were selected, because only at these locations the procedure of converting detections into counts (as described in chapter 0) apply. The second step was evaluating the quality of the data from the loop detectors at the selected intersections. Only the detectors providing a constant flow of detection data were selected. Count locations where the data showed many gaps (i.e. intervals where no data was stored) were removed from the selection. The count data of the remaining count locations was then converted into estimated volumes according to the procedure described in section 3.3. For every count location the average count volumes for the four time periods were constructed (i.e. working day, morning peak, afternoon peak and weekend day). In constructing the averages days with non-regular circumstances were removed from the data. Public holidays, school holidays and days with public events cater for lower traffic volumes on weekdays and were therefore removed. Road works at or near a count location can cause a temporal drop in the count data or an increase due to rerouting. Therefore during the time of road works nearby count locations were removed from the selection. The remaining days were used to construct the average count values for the various count locations. Additionally a quality measure was calculated to be used in the matrix estimation stage. This quality measure was based on the number of days included to construct the average count values relative to the maximum possible days of the traffic count period. These average counts and the accompanying quality measures were then imported into the RVM to be used in the matrix estimation procedure.

**Matrix estimation based on traffic counts**

Matrix estimation provides a series of techniques for estimating the travel demand from a variety of data sources such as observed link volumes, home interviews, roadside interviews, on-board public
transit surveys, etc. in this case data from traffic lights (i.e. estimates for observed link volumes) and the NTS (i.e. a nationwide home interview) were used. The first step is already conducted in the phase of estimating an initial OD-matrix (see section 4.3.1). In this phase the traffic counts are added and the matrix is re-estimated to be in line with the actual bicycle counts in the network.

The matrix estimation procedure (taken from Omnitrans) focuses on aligning the various initial matrices with the traffic counts (i.e. average volumes for the entire workday) and is executed as follows:

\[ T_{ij} = T_{ij} * X_{ij}^{\delta_k} * X_{0} \]  

Equation 4.4

Where:
- \( T_{ij} \) = number of trips from zone \( I \) to zone \( j \)
- \( X_{ij} \) = update factor for specific OD-pair based on the relevant count value
- \( \delta_k \) = count weight and measure of reliability of the constraining count value
- \( X_{0} \) = update factor for entire OD-matrix

\[ X_{ij} = \sum_c \left( \frac{R_1}{\delta_c * T_c} \right) \]

Where:
- \( R_1 \) = count value of first count location
- \( \delta_c \) = set of OD-pairs passing by count location \( R_1 \)
- \( T_c \) = trips in the OD-matrix

\[ X_{0} = \frac{\sum_{\alpha} R_{\alpha}}{\sum_{\alpha} M_{\alpha}} \]

Where:
- \( \sum_{\alpha} R_{\alpha} \) = sum of all count values
- \( \sum_{\alpha} M_{\alpha} \) = sum of all trips passing by count location in OD-matrix to be re-estimated

The matrix estimation procedure is an iterative process. The number of iterations influences the accuracy and calculation time. More iterations will increase accuracy of the bicycle load estimations but will also increase calculation time. As a start we chose to execute 5 iterations for every matrix estimation. A smaller number of iterations is possible when a small number of constraints are present. In this case we used 37 count locations.

4.4 Results

In this chapter the results of the modelling study are presented. We used Omnitrans to estimate OD-matrices and for the assignment of the trips to the network. We present and describe the variations in the results when different input variables are used and how this affects the value and usefulness of combining traffic counts with OD-pairs to come to urban bicycle flows. Basically, we varied in three aspects in the process: (1) the input, (2) the network and (3) the modelling settings. To compare the
various OD-matrices we used the trip attraction and production, the trip length distribution and the comparison with the actual counts as the indicators for the accuracy of the OD-matrices.

Firstly, the influence of the initial OD-matrix is investigated. We used three input matrices and used three more after the matrix calibration process using the traffic counts.

1. Matrix constructed with the gravity model
2. Matrix based on the National Travel Survey
3. Mix between (1) and (2)

We compared the resulting productions and attractions, trip distributions and link loads with the trip length distribution from the NTS and the actual traffic counts retrieved from the traffic lights.

4.4.1 Trip production and attraction
A first measure of comparison is the total number of trips in Enschede. According to the NTS a citizen of Enschede makes on average one bike trip per day. To make a fair comparison between the matrices the total number of trips is used to scale the matrices towards an average number of trips per inhabitant to 1 trip per workday. This implies the total number of trips produced by zones in Enschede is 155,000. The table presents the total production of bike trips for the municipality of Enschede as a whole and the four areas (city centre with the main shopping area, the north/west containing residential areas and the university, the east containing mainly residential area, and the south west containing residential areas and the main industrial area). From the table a first comparison can be made between the various initial OD-matrices before the matrix calibration procedure.

Table 4.2: comparison of initial matrices based on bike trip generation

<table>
<thead>
<tr>
<th></th>
<th>Gravity model</th>
<th>NTS</th>
<th>Mixed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of trips</td>
<td>155000</td>
<td>155000</td>
<td>155000</td>
</tr>
<tr>
<td>Trips to/from centre</td>
<td>30298</td>
<td>41093</td>
<td>41094</td>
</tr>
<tr>
<td>Trips to/from north/west</td>
<td>32110</td>
<td>33958</td>
<td>33958</td>
</tr>
<tr>
<td>Trips to/from east</td>
<td>31958</td>
<td>29072</td>
<td>29071</td>
</tr>
<tr>
<td>Trips to/from south/west</td>
<td>60634</td>
<td>50878</td>
<td>50878</td>
</tr>
</tbody>
</table>

The main difference between the matrices in terms of productions and attractions is the city centre and the south-western part of Enschede. Because for constructing the OD-matrix using the gravity model using the number of inhabitants and jobs of zones as the main variables for trip production and attraction, the trip generation in the main shopping areas in the city centre are underestimated in relation to the NTS. On the other hand the trip generation as a result of the number of jobs in the south-western part is overestimated. The share of the car in commuting to the industrial and commercial area is higher than in other parts of Enschede. It is assumed that the OD-matrix based on the NTS resembles the actual trip generation better.

4.4.2 Trip length distribution
The second measure for the comparison of OD-matrices is the trip length distribution. This measure can only be compared after assigning the various initial OD-matrices to the network. We chose to use an All-Or-Nothing assignment method with distance as the sole driver for route choice. In the
following figure we present the bicycle trip length distribution resulting from the various initial OD-matrices.

![Figure 4-8: comparison of initial matrices based on the trip length distribution](image)

The gravity model matrix resembles the trip length distribution that was directly extracted from the NTS (NTS reported), because the distribution function used in the gravity modelling was directly based on the trip length distribution of the NTS. The matrix based on the origins and destinations reported in the NTS shows higher shares of short distance trips. This can be explained by short distance shopping trips to the city centre of Enschede, which are underestimated in the gravity model matrix as a result of the underestimated trip attraction of the city centre compared with the NTS. When mixing these matrices the trips between 3 and 5km show higher shares in comparison with the gravity model and the NTS matrix. This is also caused by the shopping trips to the city centre of Enschede. As a result of the distribution function more trips from the residential areas surrounding the city centre to the shopping area are estimated. From the perspective of trip length distribution all initial matrices appear reasonable.

4.4.3 Counts
The third measure is the comparison of the initial OD-matrices with the traffic counts in the network. The loads on the network resulting from assigning the initial OD-matrices according to an All-Or-Nothing assignment solely based on distance are compared with the average daily bicycle counts on workdays at the various locations. The following figure presents the counts and the estimated loads for the three initial matrices.
Generally the three matrices overestimate the loads compared to the counts. Moreover, when evaluating the total sum of bicycles detected on a daily basis and comparing it with the estimated loads as a result of the various initial matrices, the total sum of estimated loads at the count locations is a factor 1.5 higher than the actual counts. The count at the Hengelostraat have a high share in the overestimation as 4 of the 5 count locations with the highest volumes are at the hengelostraat near the singel. These count location serve the highest bicycle volumes and are overestimated by a factor 1.5 to 2.5 and therefore account for the largest part of the overestimation. The gravity model performes best in this situation because that matrix holds less trips between Enschede and Hengelo compared with the NTS. Apparently the assignment procedure assigns all traffic to the Hengelostraat whilst in reality some cyclists may use minor roads to avoid the traffic lights and consequently also the detectors at the count locations. This means that either the network should be included in more detail (to account for the shortcuts) or the OD-matrices should be adjusted to account for bicyclists making trips but not using the modelled traffic network and that are therefore not visible at the count locations.

To test the latter the OD-matrices were rescaled according to the difference between the sum of the estimated loads at all count locations and the sum of all counts (i.e. workdays). After scaling the total sum of loads and counts at count locations is now equal. The following figure now emerges.

Figure 4-9: comparison of count values and estimated loads for the various initial matrices
After scaling with reference to the total number of daily bicycles at count locations the gravity model appears to perform best, especially for the locations with high bicycle counts. Compared with the other matrices the NTS-matrix mainly overestimates the number of trips on the Hengelosestraat.

At this point the matrix calibration procedure was executed to investigate the improvements with reference to the counts. Basically the procedure rescales the number of trips of OD-pairs assigned to a link with bicycle counts. This implies the loads will reflect the actual counts better, but it also has an impact on the total trip generation and the trip length distribution. Because this impact is small (matrix estimation reduces trip generation not more than 5%) we only focus on the differences in loads and counts. After the matrix calibration the comparison between the observed and estimated number of bicycles gives the following results.
Figure 4-11: comparison of count values and the calibrated loads for the various initial matrices

The calibration procedure produces estimated loads that are in line with the counts at the count locations. When investigating the standard deviation of the estimated loads against the actual counts the gravity model initially (i.e. before the matrix calibration) performs best (standard deviation of the difference between the estimated loads and the counts is the lowest for the gravity model). This is mainly caused by the count locations at the Hengelostraat serving high bicycle volumes and having a high share in this comparison. According to the NTS there are more trips between Enschede and Hengelo than the counts show. After the matrix calibration procedure all three matrices perform comparably.

Table 4-3: comparison of the performance of the initial OD-matrices

<table>
<thead>
<tr>
<th></th>
<th>Initial matrix</th>
<th>After scaling</th>
<th>After matrix calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravity model</td>
<td>5531</td>
<td>3307</td>
<td>361</td>
</tr>
<tr>
<td>NTS</td>
<td>7853</td>
<td>4749</td>
<td>251</td>
</tr>
<tr>
<td>Mixed model</td>
<td>6390</td>
<td>3625</td>
<td>338</td>
</tr>
</tbody>
</table>

When comparing and reviewing the bicycle flows resulting from the assignment, the following figures arise. The figures show the differences between the loads resulting from the gravity model before and after matrix calibration (left) and the NTS matrix before and after matrix calibration (right). The green lines reflect the extent to which the estimated loads are in line with the counts. A green-yellow load suggests the initial matrix overestimates the loads on the particular link (i.e. overestimation is the yellow part) and a green-red load indicates the estimated load is lower than the counted volumes.
The main difference between the two figures in the estimated loads on the Hengelosestraat. According to the gravity model there are fewer cyclists using the Hengelosestraat in comparison with the NTS matrix. However according to the trip data from the NTS there is a larger than expected relation between Hengelo and Enschede in terms of bicycle traffic. As this relation is extracted directly from the trip data it is assumed that the NTS matrix better reflects the actual situation than the gravity model where this relation is estimated. A further reason for the overestimation of estimated loads on the Hengelosestraat is the assignment method. Because the Hengelosestraat is the shortest route between Hengelo and Enschede in terms of distance many trips are assigned to this route, while the bicycle highway along the railway has a higher utility (less interruptions, more comfortable). In the next section the effects of the assignment method on the estimated loads is investigated.

4.4.4 Assignment
In the network the lengths and the expected speeds at bicycle lanes were already defined. However, in the speeds may affect the shortest paths and therefore route choice. To test the sensibility of these characteristics the speeds on the bicycle paths was varied and for resulting differences in bicycle flows to be evaluated.

All-or-nothing assignment based on various combinations of weights for time and distance. Obviously the results depend on the characteristics of the network. In the network from the RVM model the speed for bikes on all paths and roads is fixed at 15 km/h. Only the cycling highway along the railway has a different speed characteristic. Therefore the only choice for all cycling trips is whether or not to use the cycling highway. The results seem to have improved when relying on the travel time; however this result is highly dependent on the distribution of trips between the Hengelosestraat and the cycling highway. Four of the five count locations recording the highest volumes are located on the Hengelosestraat and are strongly affected by the variations in the assignment.
The figure shows that using time as the sole driver for bicycle route choice in the RVM network shifts loads from the Hengelosestraat to the bicycle highway. The estimated loads are now more in line with the counts. This implies that the bicycle highway is not used for being the shortest route, but increases in importance when travel time is taken into account. Therefore route choice in bicycle trips is an important field of study in this context. Adding GPS traces of cyclists would provide more insight in this matter. As the data is not available within the context of this report, we can manipulate the length of the road segments that form the bicycle highway to reflect the higher utility. However, we chose not to, because we cannot test the results as there are no bicycle detectors at the bicycle highway to support this manipulation.

4.5 Discussion
In this chapter the idea of combining the bicycle count data from the traffic lights with the trip data from the NTS was investigated in an effort to estimate urban bicycle flows. Bicycle count data excel in presenting dynamics in daily bicycle volumes at the specific count location, but lack the origin-destination component to estimate flows. The NTS on the other hand contains origin-destination relations, but lack the critical mass of trips to construct a full OD-matrix for bicycle traffic on the scale of Enschede. Combining the two sources may exploit the advantages of both data sources to form a new bicycle information source. From the NTS (edition 2004 to 2013) the trips in and around Enschede were selected. Aggregates such as average daily bike trip rate per inhabitant and total trip generation of the various postal zones in Enschede were extracted to get an overview of the trip production, trip attraction and distribution of bicycle traffic in Enschede. Various initial OD-matrices were proposed and assigned to a traffic network extracted from the RVM. The resulting bicycle loads
on the network were then compared with the actual counts at the count locations and a matrix calibration procedure was conducted to align the matrices with the counts.

The total trip generation of trips extracted from the NTS in combination with the number of inhabitants and jobs per postal code provides first glance at the most important bicycle trip generators. On average an inhabitant of Enschede makes 1.03 bicycle trips per day. This implies that a first estimate of the trip generation of postal zones can be made. With the results of the analysis of the trip generation and trip length distribution three initial OD-matrices were generated: (1) gravity model, (2) NTS matrix and (3) a mixture between the gravity model and the NTS matrix. These matrices were assigned to the network and compared with the bicycle count data in the network.

The initial matrix based on the gravity model presumes knowledge of the bicycle as the trip generation of the zones and the trip length distribution of bike trips should be available. These aggregate values are used to construct the second initial matrix (i.e. gravity model). This matrix relies on the trip generation based on the socio-economic contents of the centroids and the trip length distribution extracted from the NTS. The resulting matrix underestimates the number of bicycle trips generated in the city centre of Enschede and overestimates the bicycle trip generation of the industrial and commercial area in the western part of Enschede. This is caused by the absence of some essential characteristics in bicycle trip generation in the construction of the initial matrix. Shopping area and number of student places could not be included. Although the trip length distribution and traffic counts resemble the NTS trip length distribution and the actual traffic counts, the initial matrix fails to take the higher bike rate to the city centre and the lower trip rate to the main commercial area in the west into account. When more information is available about the location and size of shops and schools the trip generation component of this initial matrix can be improved to better reflect the bicycle trip production and attraction.

The second initial matrix used the trips as registered in the NTS to construct an OD-matrix. After the trip assignment on the network the results show that the trip length distribution is slightly different from the trip length distribution retrieved directly from the NTS. The NTS matrix has a higher share of short distance trips. This can be caused by inaccurate trip length estimates in the registration of trips in the NTS. The comparison with the counts in the network showed an overestimation of the bicycle loads in general, primarily caused by a high estimated load on the count locations at the Hengelosestraat. In this measure the NTS matrix is outperformed by the gravity model matrix. Still, we believe that the NTS matrix resembles the bicycle flows in Enschede better, however the flows aren’t assigned to the proper routes. For example, the bicycle highway should serve as one of the main routes between Enschede and Hengelo, but this is not at all reflected by the assignment procedure as only travel distance is taken into account. Moreover, bicycle flows do not necessarily stay within the designed space of the main arterials as is the case for cars. A cyclist can avoid the main arterials and therefore the count locations more easily.

The calibration of the various matrices to the bicycle counts catered for a situation where all matrices resemble the actual traffic counts. The NTS matrix and the gravity model perform similarly. However, because the NTS matrix is assumed to better reflect the specific OD-relations in Enschede the NTS matrix is used in the rest of this study.

In this study we showed it is possible to estimate bicycle flows based on an OD-matrix extracted from the NTS and to align it with the actual bicycle counts in the network. This implies a new information
source for urban bicycle traffic can be generated by combining trip data from the NTS and bicycle volumes from count locations in the network. Also when less specific data than in the NTS is available (i.e. solely based on trip length distribution and trip ends), still a reasonable estimation of bicycle flows can be produced. This allows for presenting and studying urban bicycle flows and its dynamics. Further specification is possible as well. For example, when the initial OD-matrix is calibrated according to the count values of a specific year, season, week, weekday or peak period, the bicycle flows for specific time periods can be generated and comparisons of consecutive years, weeks or days can be made. This allows for a very relevant and valuable information source for bicycle policy making and evaluation.
Part II: Database and application

The second part of the project is concerned with the computer science component of the project and focusses on the storage of data and presentation of the bicycle traffic information. In this part of the report the data as presented in the previous chapters is structured and stored in a database. In chapter 5 the database design and implementation are described that hold the bicycle related data for Enschede. In chapter 6 the design and implementation of an application is described and the final design in presented. The seventh and final chapter provides directions for further research and development of bicycle monitoring in the urban environment in general and for the case Enschede in particular. Moreover, further suggestions for the database and application development are discussed.
5 Database design

This chapter presents the process of designing a database to hold the bicycle traffic information as described in the previous part of this report. The database design is strongly related with the application design that is described in chapter 6.

5.1 Introduction

The design of a database structure starts with a clear and widely accepted description of the problem or situation the database will be used for. The basic concepts and functionalities must be defined to describe the structure, the operations and the constraints the database is subjected to. The purpose of a database is “to provide a reliable repository for a potentially large volume of well-structured data to a large community of end-users, who may want to use the repository simultaneously under acceptable performance requirements” (Morales, 2014). To get to a database design a relational data model can be constructed. Such a model specifies all data and presents the interrelation between the data classes. The process of designing and maintaining a database can be captured in the following steps (Morales, 2014):

- System definition
- Requirements collection and analysis
- Database design
- Application design
- Prototyping (optional)
- Implementation
- Testing
- Operational maintenance

In this chapter the process of designing a database for bicycle flows in Enschede is discussed following the aforementioned steps.

5.2 System definition

The first step of system definition describes scope and boundaries of database system and the major user views. In this case the definition follows the problem definition defined in the introduction of this report. In summary the problem is defined as follows:

The municipality of Enschede invests in maintaining and improving accessibility of the urban transport system. Especially the accessibility for bicyclists will be supported by the municipality, not only to improve the situation for the cyclists, but also to relieve the congestion problems for car traffic. Currently many municipal investments in transport go to measures to improve infrastructure for cycling. However, the municipality lacks the information from the network to support and evaluate the decisions for investments in cycling infrastructure, essential for sound policy making.

On the other hand, the municipality (as road authority) already collects traffic data to a large extent, but does not use it to its full potential. There is a surplus of sensors in the traffic network mainly near traffic lights to serve the traffic light control system. This implies the municipality of Enschede has a rich source of traffic data. Although the data in principal holds information about traffic volumes for both vehicles and bicycles, it is not used for that purpose.
Secondly there is data available from the National Travel Survey. This is a nation-wide survey, conducted on a yearly basis, to collect information about trips people make in the Netherlands. This survey also contains trips within the municipality of Enschede. Information about the number of trips and the modal split of inhabitants can be retrieved.

Using the data that is already collected and presenting the information from the various sources can already help in the provision of policy relevant traffic information. However, combining the different data sources may provide richer information, because it may bring the advantages of the different sources together.

5.2.1 Objective
The objective is to construct a database to support an application presenting the bicycle volumes and flows in the city of Enschede for the municipality to get more insight in the bicycle volumes and flows in the city (e.g. volume profiles over a specified time period at a selected count location, estimated origins and destinations of cyclists at a selected count location, estimated loads on the network at locations without a bicycle counter) and can provide a central role in the bicycle traffic monitoring.

5.2.2 Potential users
In the design it is important to realize who may be the potential end-users and what are their specific needs and typical usage of the database and the application. Primarily the database and the resulting application are designed for the municipality of Enschede to view and analyse the available traffic data and the processed traffic information for bicycle flows within the municipal borders of Enschede. The needs are derived from discussions with the municipality as primary user. For the municipality the main functionalities are listed:

- Central database for bicycle data (traffic lights, NTS, manual counts)
- Control, maintain and update the database
- Overview of bicycle volumes on main bicycle arterials
- Volumes and their dynamics of specific arterials
- Overview of travel behaviour in Enschede
- Overview of estimated bicycle flows (origins and destinations)

5.2.3 Functional requirements
The potential users and their objectives result in the following functional requirements:

- Allow insight in the traffic related data collected and available for the municipality
- Combine data sources to get enriched information about bicycle traffic
- Create policy relevant information about the bicycle traffic
- Provide a means to present the information and to enable analysis of the information

5.3 Data collection and analysis
5.3.1 Description of data
As input for the database there are basically three sources of data, being: (1) traffic light data, (2) data from the national travel survey and (3) the traffic network.

Traffic light data
The database should be able to process two levels of traffic light data. For the application data about traffic volumes for a specific time interval are required. This implies the basic traffic light data needs
to be processed into volumes per interval. A sample of the basic traffic light data is presented in Table 5-1.

**Table 5-1: sample of traffic light data**

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Duration</th>
<th>Location</th>
<th>Info</th>
<th>Type</th>
<th>ID</th>
<th>ID_COND</th>
<th>Q</th>
<th>Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D02_1</td>
<td>-</td>
<td>100</td>
<td>7</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D02_2</td>
<td>-</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D03_1</td>
<td>-</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D03_2</td>
<td>-</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D05_1</td>
<td>-</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D05_2</td>
<td>-</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D06_1</td>
<td>-</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D06_2</td>
<td>-</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D08_1</td>
<td>-</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D08_2</td>
<td>-</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D09_1</td>
<td>-</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D11_1</td>
<td>-</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D11_2</td>
<td>-</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D11_3</td>
<td>-</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D11_4</td>
<td>-</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D11_5</td>
<td>-</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D12_1</td>
<td>-</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D12_2</td>
<td>-</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D22_1</td>
<td>-</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D24_1</td>
<td>-</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D26_1</td>
<td>-</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>1-1-2010</td>
<td>0:00:00</td>
<td>900</td>
<td>ENSHAGE</td>
<td>COUNT_OCCUPIED</td>
<td>DETECTOR</td>
<td>D28_1</td>
<td>-</td>
<td>100</td>
<td>1</td>
</tr>
</tbody>
</table>

The figure shows the count data from one intersection for one time interval of 15 minutes. The data consists of the timestamp (date and time), the duration of the interval (900 seconds), the nametag of the intersection, the type of data (count of detections), the detector type and name and the actual count and a quality measure for the count value. As was described in section 2.1 this data needs to be processed and converted into estimations of bicycle volumes. After this processing step for every branch of the signalized intersection (i.e. corresponding to one link in the network) the estimated bicycle volume per time interval is stored. In the following table this bicycle volumes for a single branch is presented for eight consecutive time intervals.

**Table 5-2: sample of processed traffic light data**

<table>
<thead>
<tr>
<th>TimeInterval</th>
<th>drukteXP</th>
<th>intensiteit</th>
<th>Qintensiteit</th>
<th>Qinterval</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-01-01T00:00:00</td>
<td>1</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2010-01-01T00:15:00</td>
<td>5</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2010-01-01T00:30:00</td>
<td>17</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2010-01-01T00:45:00</td>
<td>24</td>
<td>5</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2010-01-01T01:00:00</td>
<td>34</td>
<td>10</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
The processed data consists of a time interval, a value for the total number of bikes passing during the time interval for the entire intersection, the number of cyclists counted in the direction of interest, a quality measure of the counted cyclists (i.e. detection pattern of the particular detection loop) and a quality measure of the data collection at the intersection of interest (i.e. all detection loops functioned as expected).

The processed data is used as input for the application. Therefore the database must contain the processed data. However, when new data needs to be introduced in the database the processed data must be constructed, either from the basic data or in the shape of the processed data. As a first step the database will be based on the processed data.

**NTS data**

The data from the National Travel Survey consists of trips of the Dutch population. A more elaborate description of the data source was given in section 2.2. The data reflects the trips an individual makes on one particular day. One record reflects one trip with its associated personal, household and trip characteristics. These characteristics include the gender and age of the participant (traveller), the household composition, income and availability of travel modes and the characteristics of the trip such as the timing (year, month, day, time of day and travel time), origin, destination, distance and travel mode. The following table shows a small slice of the data.

**Table 5-3: sample of NTS data**

<table>
<thead>
<tr>
<th>OP</th>
<th>OPID</th>
<th>Steekproef</th>
<th>Mode</th>
<th>HHPers</th>
<th>Jaar</th>
<th>Maand</th>
<th>Dag</th>
<th>VertPC</th>
<th>AankPC</th>
<th>AfstV</th>
<th>Hvm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>9711</td>
<td>1079</td>
<td>2022</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>9711</td>
<td>1079</td>
<td>2022</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>9711</td>
<td>1079</td>
<td>2022</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>1079</td>
<td>3524</td>
<td>502</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>1079</td>
<td>3524</td>
<td>502</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>1079</td>
<td>3524</td>
<td>502</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>1079</td>
<td>3524</td>
<td>502</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>3524</td>
<td>9711</td>
<td>2008</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>3524</td>
<td>9711</td>
<td>2008</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>12301039</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2012</td>
<td>1</td>
<td>1</td>
<td>7827</td>
<td>7821</td>
<td>50</td>
<td>6</td>
</tr>
</tbody>
</table>

In the case of bicycle flows within the municipality of Enschede a selection of trips and variables reduced the amount of data to be processed. To get the relevant data from the NTS the consecutive editions need to be combined, the appropriate trip characteristics need to be selected and the relevant trip origins and destinations need to be selected. The NTS editions differ in the naming of the variables from year to year. This requires a processing step to map the same variables of different years on each other. From the NTS only a limited number of variables is relevant in this context (i.e. origin and destination of a trip, trip modality, trip timing (date, time), trip duration and distance, trip purpose and trip weight factor). The other variables are therefore filtered out. The last
processing step is filtering based on origins and destinations and selecting only the trips in and around Enschede. Assuming that bicycle traffic mainly concerns trips shorter than 15 km, trips from and to locations outside Enschede and its direct surroundings were filtered out.

**Traffic network**

The traffic network used in the database was retrieved from a traffic model used in the Region Twente. This network was described more elaborately in section 2.3. The network as it was loaded into the database consisted of centroids (the origins and destinations of trips), nodes and links (road segments connecting the centroids and nodes). Moreover the locations of the detectors near traffic lights are included in the network. In the database the following tables were imported.

**Table 5-4: sample of centroids**

<table>
<thead>
<tr>
<th>CENTROIDNR</th>
<th>NEWNR</th>
<th>NAME</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000001</td>
<td>7511</td>
<td>258095</td>
<td>471345</td>
</tr>
<tr>
<td>2</td>
<td>1000002</td>
<td>7511</td>
<td>258277.5</td>
<td>471387.8</td>
</tr>
<tr>
<td>3</td>
<td>1000003</td>
<td>7511</td>
<td>258274.6</td>
<td>471128.1</td>
</tr>
<tr>
<td>4</td>
<td>1000004</td>
<td>7511</td>
<td>258051.9</td>
<td>471096.7</td>
</tr>
<tr>
<td>5</td>
<td>1000005</td>
<td>7511</td>
<td>257834.7</td>
<td>471104.8</td>
</tr>
<tr>
<td>6</td>
<td>1000006</td>
<td>7511</td>
<td>257625</td>
<td>471295</td>
</tr>
<tr>
<td>7</td>
<td>1000007</td>
<td>7511</td>
<td>257829.4</td>
<td>471316.1</td>
</tr>
<tr>
<td>8</td>
<td>1000008</td>
<td>7511</td>
<td>257570.7</td>
<td>471441.2</td>
</tr>
<tr>
<td>9</td>
<td>1000009</td>
<td>7511</td>
<td>257758.5</td>
<td>471466.3</td>
</tr>
<tr>
<td>10</td>
<td>1000010</td>
<td>7511</td>
<td>257879.5</td>
<td>471483</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**Table 5-5: sample of nodes**

<table>
<thead>
<tr>
<th>NODENR</th>
<th>NAME</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>191264.00000000</td>
<td>508141.00000000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>138977.69789086</td>
<td>483849.44625617</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>161527.00000000</td>
<td>383787.00000000</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>106546.56308080</td>
<td>447909.43050027</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>255903.62500000</td>
<td>484660.68750000</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>207256.00000000</td>
<td>372041.00000000</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>256632.26562500</td>
<td>484148.34375000</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>256863.95312500</td>
<td>484109.46875000</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>233981.56250000</td>
<td>470147.50000000</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>234168.59375000</td>
<td>470307.71875000</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**Table 5-6: sample of links**

<table>
<thead>
<tr>
<th>LINKNR</th>
<th>NAME</th>
<th>DIRECTION</th>
<th>LENGTH</th>
<th>ANODE</th>
<th>BNODE</th>
<th>TOEDELINAB</th>
<th>TOEDELINBA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0.93000001</td>
<td>1003151</td>
<td>4310</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.28999999</td>
<td>1003152</td>
<td>3809</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.36000001</td>
<td>1003152</td>
<td>117943</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.66000003</td>
<td>1003153</td>
<td>4347</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
</tbody>
</table>
In the database each link was converted into two links when the specific link is two-directional. In this way, the estimated loads can be assigned to one specific link and direction instead of having two loads per link.

Table 5-7: sample of count locations

<table>
<thead>
<tr>
<th>COUNTNR</th>
<th>NAME</th>
<th>LINKNR</th>
<th>DIRECTION</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>866</td>
<td>Varviksingel(W)</td>
<td>1186</td>
<td>2</td>
<td>257987.2</td>
<td>470107.2</td>
</tr>
<tr>
<td>868</td>
<td>Kuipersdijk (N)</td>
<td>1182</td>
<td>1</td>
<td>258032</td>
<td>470165</td>
</tr>
<tr>
<td>869</td>
<td>Varviksingel(O)</td>
<td>1184</td>
<td>2</td>
<td>258081.8</td>
<td>470122.9</td>
</tr>
<tr>
<td>871</td>
<td>Kuipersdijk (Z)</td>
<td>1187</td>
<td>2</td>
<td>258076.6</td>
<td>469968.7</td>
</tr>
<tr>
<td>873</td>
<td>Getfertsingel(W)</td>
<td>1340</td>
<td>1</td>
<td>257067.4</td>
<td>470260.1</td>
</tr>
<tr>
<td>874</td>
<td>Getfertsingel(W)</td>
<td>1340</td>
<td>2</td>
<td>257067.4</td>
<td>470260.1</td>
</tr>
<tr>
<td>876</td>
<td>Broekheurnerweg</td>
<td>1341</td>
<td>2</td>
<td>257224.7</td>
<td>470354.3</td>
</tr>
<tr>
<td>877</td>
<td>Getfertsingel(O)</td>
<td>1338</td>
<td>1</td>
<td>257240.2</td>
<td>470236.9</td>
</tr>
<tr>
<td>879</td>
<td>Burg.v.Veenlaan</td>
<td>1339</td>
<td>2</td>
<td>257122.8</td>
<td>470192.9</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

5.3.2 Use of data

Traffic light data

The traffic light data is used to generate a profile of the bicycle intensities during an average day. For a workday, morning peak, afternoon peak and a weekend these profiles will be generated for each count location. Figure 4-2 shows examples of these profiles.

Moreover, the average count values will be used for the calibration of the estimated bicycle flows. Based on the NTS data and origin-destination matrix is constructed. After assigning the trips per OD-pair to the network the resulting assignment will be calibrated according to the count values at the count location.

NTS data

The NTS data is used to generate the total production and attraction of trips. For each postal zone in the municipality of Enschede and its direct surroundings the total number of trips originating from and arriving at the specific postal zone is calculated. Moreover the modal split of departing and arriving trips is presented.

Figure 5-1: example of modal split of zone
This information can be presented on a yearly basis to enable the investigation of year-to-year variability of the production, attraction and modal split.

The interrelations between postal zones are also retrieved from the NTS data. An origin-destination matrix is constructed, representing the traffic flows in the network. This OD-matrix will be assigned to the network to estimate bicycle loads on the bicycle paths in the network.

Traffic network

The traffic network is used as a carrier of the bicycle information that is available. The bicycle counts are collected at a certain location in the network, trips from A to B use the network to move and the total sum of trips is projected on the network as loads on the individual links. To fulfil these requirements the shortest routes between all origins and destinations need to be calculated. All OD-pairs now have a set of links to be used to get from the origin to the destination. Moreover, all links now have a number of OD-pairs that make use of that specific link. The routes are also stored in the database. The resulting loads on the links can be represented similar to the following figure.

Figure 5-2: example of traffic network as information carrier

5.4 Database design

Based on the description of the ‘problem’ a database structure can be designed according to the following steps: (1) functional design, (2) conceptual database design and (3) physical database design. The first step is concerned with constructing a data model of the functions and the relations leading up to the availability of the requested bicycle traffic information. It describes the types of data and data processing that occurs in a database application and is in line with the user requirements. It results in a high-level definition of the information contents of the database. The second step maps the functional design into a definition of classes and relations for a relational database (i.e. class diagram). The result is the conceptual database schema. In the third step the database is actually implemented in accordance with the previous designs. In this section we focus on the first two steps. The third step will be mentioned in section 5.5.
As mentioned before the concept of the database is to store the bicycle traffic related data that enables an application to present bicycle traffic information. The information therefore needs to be connected to the bicycle network. Figure 5-3 shows the general structure of the database.

Basically the bicycle network consists of centroids (the origins and destinations on trips), links (i.e. the roads) and counts (i.e. locations in the network where bicycle counts are available). Obviously the bicycle volume data is related to the count locations. The count location is related to a road segment. From the trip data an OD-matrix can be composed, holding the number of trips between the centroids in the bicycle network. The links in the network now determine the possible route between the origin and destination centroid of interest. With this functional design the data sources are linked to the network being the information containers and the required information can be extracted from this design. The bicycle volume profiles can be calculated per count location. The loads on the network can be estimated by the summation of the trips of the OD-pairs using the links in their route.

The aforementioned functional design is formalized into a conceptual representation of the database and is presented in the following figure. It can be considered as a model-based design approach in the sense that it establishes a database framework with functional characteristics independent from the deployment.
Figure 5-4: conceptual database design (i.e. class diagram)

Definitions:

A **Link** is a road segment between two nodes with a direction (from node A to node B or the other way around). A link has characteristics such as travel modes using the link, maximum speed and capacity.

A **Node** is a point in space connecting links. Every node connects at least two links (possibly more) and can be viewed as an intersection.
A **Centroid** is a specific node (point in space) containing the socio-economic content of an area. A centroid represents an area in terms of number of inhabitants, number of jobs, etc.

A **Count** is a location where detectors are present in the traffic network providing actual traffic volume data to the database. A count is always related to one specific link.

A **Route** is a set of links representing the route between two centroids (origin and destination). Between a specific OD-pair there can be multiple routes. Moreover a route can be multimodal accounting for multimodal trips.

An **OD** pair is a combination of two centroid representing an origin and destination of a trip.

**TrafficLightData** consists of the data from the detectors at the traffic lights (in Enschede). The data can come from a loop detector or a signal detector that detect the number of vehicles passing and the number of cycles (count of the green phases) over a certain time duration.

**TripData** consists of data from the National Travel Survey in the Netherlands, containing trips in the Netherlands with their accompanying trip characteristics (e.g. timing, origin, destination, travel mode) and personal characteristics (e.g. age, gender, income of the trip maker).

An **Irregularity** is a non-regular situation with a time and location component. For example, a public holiday is an irregularity that affects traffic on that particular day, while road works may affect the loads on a specific link (location) for a short time. These irregularities are important for the calculations of average values for regular traffic volumes.

A **PostalZone** is an area based on the system of assigning identifiers to addresses and areas. These zones generally are bound by natural borders (e.g. water ways and roads) and reflect a distinct neighbourhood.

### 5.5 Implementation

In the next section the implementation of the aforementioned database design is described. A pragmatic approach was adopted of firstly loading the processed data tables in the database and secondly implemented the dependencies and relations between the tables.

#### 5.5.1 Loading data

In the database and the application processed data is used. For all count locations the estimated count values for all 15-minute intervals in 2010, 2011 and 2012 are loaded into one table. Also a pre-processed table containing the relevant personal and trip characteristics of trips in and around Enschede was directly loaded into the database. Loading the traffic network required more effort. The network consists of multiple tables; centroids, links, nodes and count locations. To avoid confusion between the centroids and the nodes the numbering of the centroids was altered to prevent overlap in the identifiers of centroids and nodes. Otherwise ambiguity could occur about the start and end point of a link. Moreover the table containing the routes is constructed beforehand. By running a shortest route algorithm from all origins to all destinations and storing the links used in the shortest routes the routes table is constructed and loaded.
5.5.2 Processing data and implementing data structure

After the loading of the data, the types of the data in the database are all set to the default type, being a text string. In order to enable querying and do calculations on the data, the data types should be set to their appropriate types. This was done using sql-commands on the specific tables in the database. For example, the table containing the count value data from the count locations at traffic lights needs to be altered to enable the calculation of average numbers of bicycles and daily count profiles.

```sql
ALTER TABLE aggr_countdata
  OWNER TO postgres,
  alter column volumexp type float using volumexp::float,
  alter column bikeint type float using bikeint::float,
  alter column qbikeint type float using qbikeint::float,
  alter column qinterval type float using qinterval::float,
  alter column datum type date using to_date(datum, 'YYYY-MM-DD'),
  alter column tijd type time using to_timestamp(tijd, 'HH24-MI-SS'),
  foreign key (countlocation) reference countconfig(countlocation) on delete cascade on update cascade;
```

Moreover to ensure consistency of the database relations and constraints should be introduced. As the database is constructed from a collection of imported tables, the database structure (as presented in Figure 5-4) still needs to be effectuated by denouncing relations and constraints on the imported tables. This is done by creating foreign keys referring to other tables in the database. Another processing operation conducted in the database is creating views. A view is a stored database query that behaves as a data table. The OD-matrix is created as a view to enable quick query of OD-relations for the estimation of loads on the links in the network.

5.6 Operational maintenance

After the preceding steps the database is fully functional. It contains the available bicycle data and the referencing is installed to ensure consistency. However, during the operational phase the existing data sources may provide recent data and new data sources relevant in the context of urban cycling may become available. In these cases some actions are required to update the database to hold this new data and to update the associated bicycle information. Currently the database isn’t designed to automatically load new data. Updates of existing data sources are added on a yearly basis. This timeframe is chosen as the NTS provides data on a yearly basis as well. Although the traffic light data provides a continuous source allowing a (nearly) real-time feed of data to the database, the objective of providing an overview of bicycle information for the purpose of policy making and evaluation does not require real-time updates of the database. Therefore also the bicycle volumes can be updated with recent data on a yearly basis.

The operational maintenance of the traffic network on the other hand may carry some problems. Currently the processing of bicycle information assumes an unaltered network. For example, the shortest routes through the network may change as a result of new infrastructure or traffic constraints. However the current database does hold the traffic network accounting for future infrastructural projects. Moreover, temporal changes in the network (e.g. due to road works) are beyond the scope of the application. Major changes in the bicycle network may demand adjustments in the database and an update in the estimated bicycle flows. This may mean the comparisons of bicycle information between consecutive years become meaningless.
5.7 Discussion

In the design and implementation a pragmatic approach was chosen. The main objective is to store the available data sources. In a way that allows the application to take full advantage of the data. The processed data (i.e. OD-matrix, routes, assignment and calibration) is generated outside of the database and imported as tables. Because the database is not designed to have a real-time feed of data, the database can be considered as ‘static’. On a yearly basis new data can be imported that may also require the database to update the bicycle information (e.g. bicycle loads and average bicycle counts). Future improvements in the database could include a more frequent feed of count data and a trigger in the database to recalculate average counts, the OD-matrix and routes on a regular basis. This will also enable a regular update of estimated loads on the bicycle network.
6 Application design

The application design and implementation is the final step in the development of the MOVE monitoring tool. The application design is an iterative process of balancing the user requirements of the municipality of Enschede and the available data and information.

6.1 Introduction

Web services could serve as a portal to get an overview of the current efforts in traffic monitoring in combining the data and information from the various traffic monitoring efforts (i.e. traffic lights and their detectors, manual counts of traffic volumes, travel surveys). The main purpose of the application is to present the available bicycle information to the user. The use of modern web technology and programming packages makes it possible to construct a highly interactive and user friendly application. The presentation should be highly visual to generate a comprehensible overview and should use maps of the respective area with the requested bicycle information as overlays and graphical representations of the bicycle data in the form of charts. Several synchronized views allow users to have multiple perspectives on the data, allowing for more in-depth understanding of the bicycle patterns in specific areas of the city. Moreover, it could serve as an input portal for manual information in combination with the current automated monitoring efforts. Currently the application uses pre-process data, but since it is based on the use of web services it allows the handling and presentation of both stored and real-time data.

The objective is to construct an application presenting the bicycle volumes and flows in the city of Enschede for the municipality to get more insight in the bicycle volumes and flows in the city (e.g. volume profiles over a specified time period at a selected count location, estimated origins and destinations of cyclists at a selected count location, estimated loads on the network at locations without a bicycle counter) and can provide a central role in the bicycle traffic monitoring. Additionally the application could serve as a communication tool towards the general public to show the municipal efforts in improving the urban transport system and for the general public to use the overview of the bicycle information. This is beyond the scope of the current study.

6.2 Application structure

The approach to display the information in an application can range from a design relying on the client and relying on the web server. Figure 6-1 shows how this can be portrayed.
In case of a thin client the application is mainly developed on a server. The storage of data and the processing to information is conducted on a server. The user can access the information by accessing the application via the web browser which in turn will send requests to the server and will present the information. On the other side of the range one can choose for a thick client. This implies that most of the applications structure is stored and processed at the client-side. Some information may be retrieved from the web, but most information is handled and presented at the client. In the next sections both a client-side application (i.e. desktop application) and a server-side application (i.e. web application) will be discussed. The primary focus in this study is to develop a desktop application first. In a later stage this can be extended to a web application.

A desktop application relies on the functionalities of the client, also for storing and processing data. In this case there are two main sources of information: (1) the database containing the bicycle related data and information and (2) web services providing referential information such as maps to project the bicycle information on. All processing steps such as retrieving data from the database, retrieving maps and processing and projecting the information onto the map is conducted on the client-side.

![Desktop Application Design](image)

**Figure 6-2: desktop application design**

This structure allows the user to develop the application and adjust it to their personal preferences. This structure is preferable in case of a small number of users and if adaptability and flexibility is required. As the application is primarily developed for the municipality of Enschede (one user) and the initial version of the application should be flexible for further ad-hoc development a desktop application is best suitable in this case. In a later stage when more users use the application simultaneously a web application is more suitable. In that case the users can simply use the web browser to navigate to the application and view the requested information. Any updates of the
application or data does not require any software updates at the client-side. However, in this development phase this has a lower priority. The initial structure of the web application and conceptual design in presented in Appendix A.

6.3 Users and functionalities
The main functionalities the application are composed from an investigation of the potential users and their information needs. The application will have two main users. The primary user is the municipality. As the municipality can use the application as a communication tool towards the general public, a second type of users appears. Functionalities for the general public can be added in a later stage and are outside the scope of this study. For the municipality the main information needs and functionalities are listed.

- Central database for bicycle data (traffic lights, NTS, manual counts)
- Control, maintain and update the database
- Overview of bicycle volumes on main bicycle arterials
- Volumes and its dynamics of specific arterials
- Overview of estimated bicycle flows (origins and destinations)
- Communication tools towards general public for bicycle related topics

The resulting application will entail four main functionalities to fulfil the requirements of the end-user of the application:

1. Traffic counts and their resulting traffic profiles
2. Traffic flows and the resulting traffic loads on the network
3. Traffic generation and modal split of neighbourhoods
4. Manage and update data sources and information

These four functionalities will be addressed in the conceptual design of the application

6.4 Conceptual design
To support the four main tasks of the application an interface with four panels or tabs is used. The traffic counts, travel behaviour and traffic flows all require a map for the spatial reference of the information presented. Moreover a selection panel needs to be present to enable the user to filtering by relevant criteria and build a detailed analysis of the specific selection. Finally the tabs need a results panel to present graphs and tables with the selected information. The conceptual design of the user interface is presented in Figure 6-3.
Figure 6-3: conceptual design of the user interface of the desktop application

The map panel will be the most prominent in the application. On the left hand side the selection panel and results panel are placed. The four buttons in the upper left part of the interface allow for the selection of the tabs to conduct the required analysis type. In the next sections the design of the distinctive tabs will be discussed in more detail.

6.4.1 Traffic counts
For the functionality of presenting the bicycle volumes the following requirements are formulated:

- Provide an overview of the traffic count locations in Enschede (Show links with a counter)
- Show traffic profiles of selected count location according to user-defined time period (this relates to the bicycle profiles as presented in Figure 4-2)
- Make a comparison of two time periods at one count location
- Add visual counts for a selected location

The following figure presents an example of the set-up of such functionality.
Figure 6-4: example of the traffic count functionality

The example in Figure 6-4 shows the daily bicycle profiles of one of the major bicycle count locations for both an average workday and a weekend day. The user may also select a specific year of interest and whether or not road works need to be filtered out of the average profiles. The map panel shows the selected count location with a red marker (instead of the regular blue markers). The results panel shows both bicycle profiles in one figure.

6.4.2 Travel behaviour in neighbourhoods

For the functionality of presenting travel behaviour (in this case trip generation, trip distribution and modal split in neighbourhoods the following requirements are formulated:

- Show the postal zones
- Allow the selection of postal zone
- Show trip generation and modal split of selected zone
- Allow for making comparison of selected zones
Figure 6-5 shows the example of how the travel behaviour is presented in the application. The user can select the area of interest, the type of information (trip generation, distribution or modal split) and the year. The results panel shows the requested table (in this case the average trip rate over the years). The map panel shows the postal zones and projects the origin-destination pairs retrieved from the MON/OViN on the map. This provides insight in the main OD-relations for the selected area.

6.4.3 Traffic flows
For the functionality of presenting the bicycle flows the following requirements are formulated:

- Show origins and destinations based on the NTS data (not assigned to the network) (using either the centroids as ODs or the corresponding postal zones)
- Show traffic loads on link based on the NTS data and calibrated with the count data. Also show the count locations that were used for the calibration
- Allow for comparison between two time periods (i.e. years) based on the calibration of the OD-matrix on the average count values for the respective years

The following figure presents an example of the set-up of such functionality.
The example above shows the bicycle flow tab of the application. The user selection part of the tab consists of two stages. First the user can construct an OD-matrix and show a part of the resulting matrix for review. In the second stage the user can select the settings for the assignment of the OD-matrix on the network. For example, the user can select the trip purpose, year, network and assignment method for the bicycle flows to be presented in the map panel. In this stage of development the choices in settings are limited as there is one bicycle network present and one assignment method available. Future versions may allow for more options.

### 6.4.4 Data management

To help in updating the data and managing the information a tab to process the traffic light data and the MON/OViN data into the relevant information is added. This information is then stored in the database and can be disclosed and presented using one of the other tabs in the desktop application. For the functionality of managing, updating and loading new data the following requirements are formulated:

- Allow the user to select new traffic light data to import
- Allow the user to select new travel behaviour data to import
- Automatically process and store the data in the database
The last tab of the application allows the user to add data to the application. The application allows the user to select a file to load into the database. While loading the traffic light data (structured as presented in Table 5-1) is processed into the information (as presented in Table 5-2) and is added to the respective database table. The travel survey data is already processed. The procedure here ensures the data labels of the consecutive years is aligned. In this way one trip data table is maintained accumulating the consecutive years of travel behaviour data.

6.5 Discussion
The desktop application is designed to be the first application to use relevant bicycle counts and flows in the transport policy process. It needs to allow the user to work with the available data sources and to conduct some basic analyses. The aim is to have a fully functional demo-version of the application. In the process of using the application new requirements and opportunities may arise to be adopted in the application. Therefore the first version of the application is to be used by the municipality of Enschede.
7 Directions for further research
The database and application reached a basic level of functionality. It offers insight in the available bicycle traffic information and allows the user to conduct basic analyses regarding bicycle volumes and flows. However, a system as such can be extended with new functionalities or updated to the state of art.

7.1 Improving on assumptions
In processing the traffic light data and MON/OViN trip data several assumptions were done to create the bicycle volumes and flows: (1) the method of converting detections at inductive loops is applicable to all inductive loops, (2) the average trip rate of citizens of Enschede extracted from MON/OViN can be extrapolated to the entire municipality and (3) cyclists choose the shortest path.

7.1.1 Processing traffic light data
The method of estimating bicycle volumes relies on 15-minutes aggregates of inductive loop detections. A more accurate method for detecting (individual) cyclists can be developed. By studying the actual distortion of the magnetic field created by the inductive loop may prove to generate a better estimate of the actual number of cyclists passing a particular point. Moreover, the configuration and placement of the loop detectors at intersections is assumed to have no effect on the detection patterns. In fact, slightly wider or longer loops or loops placed closer or further from the stop line may require other correction factors for an accurate estimation of bicycle volumes.

7.1.2 Initial OD-matrix
To be able to construct an OD-matrix for Enschede the average daily bicycle trip rate of citizens of Enschede was calculated. This number was used to raise the OD-matrix extracted from the MON/OViN to match the expected number of bike trips on a municipal level. The number of inhabitants and jobs per zone was used as a weight factor. A more accurate bicycle trip generation could be constructed using more specific characteristics of origins and destinations. For example including the number of school places in the weight factor of a zone may increase the accuracy of the bicycle trip generation as children between 6 and 17 years of age have a higher daily bike trip rate mainly caused by school trips. Also other locations generating bike trips can be included (e.g. recreational areas, supermarkets and shopping areas).

7.1.3 Bicycle route choice
Currently the route choice is assumed to be all-or-nothing on the shortest route in distance. Literature on bicycle route choice suggests that not only the distance determines the bicycle route. Broach et al. (2012) found that the chosen route is on average 11% longer than the shortest route. Other aspects such as comfort of the route, travel time unreliability (induced by signalized and unsignalized intersections) and traffic volumes also play a role in bicycle route choice (Bovy and Stern, 1990, Van Ginkel, 2014). However, in the available network these characteristics aren’t available as link attributes and are left out of the equation. Therefore, relying on the distance as main driver for route choice may seem a proper first estimate as the speed of cyclists is not limited by speed limits and therefore the shortest route is also the fastest route. A more sophisticated route choice model may improve the estimated bike loads in the network. For example interruptions, obstacles and comfort may influence cyclists in their route choice. To be able to incorporate these factors in the trip assignment procedure first more research is needed in bike route choice. Bicycle route choice is currently a major topic of research in the field of bicycle traffic. In the SMART project
in Enschede trips of people carrying their smartphone with a dedicated application installed, are registered and the data entails characteristics of their trips. The developments in using the smartphone in traffic data collection may offer new opportunities in this field as well. The data collected reflects the actual trips with their routes individuals make and may shed a new light on the choices individuals make in their cycling route. The data as such may be used directly as route choice data between origins and destinations for the selection of trips that are made in Enschede and indirectly after using the data to investigate and model route choice behaviour of cyclists. This will improve the route choice component of this study. On the other hand, the modelled network needs to have information on the route choice factors as attributes. Currently the RVM network is used as the carrier of bicycle information. The bike links in the network only contain information about the length and allowed speed of the link. A more elaborate bike network needs to be introduced to account for these route choice factors (e.g. the bike network from the Fietsersbond may be an improvement).

### 7.2 New data sources and functionalities

#### 7.2.1 Location traces and bicycle routing

In this case the available data sources shaped the design. Newly available data sources may allow new functionalities and analyses. A first major improvement may result from the inclusion of location traces of travellers. In several projects, such as i-Zone\(^2\), SUNSET\(^3\), SMART\(^4\), EMPOWER and Mobile Mobility Panel\(^5\), technologies are developed to automatically detect and register trips via a dedicated application on smartphones. This may imply that in the future more trips can be collected than is currently done in the NTS. Moreover, the data may provide more insight in the route choice of cyclists. To date this topic received little attention and is currently also disregarded in the database and application. The inclusion of this data source may require an update of both the database and the application. This implies the functional design and the conceptual database design need to be updated. Moreover, new information can be created resulting from the confrontation of the current with the new data source. The location traces may produce an extension to the route table. A particular link in the route table can then receive a portion of the trips between a particular OD-pair using the link in their route choice. The implications for the quality and functionality of the application are that the bicycle assignment can be based on actual measurements from the transport system. Moreover, the data may shed a new light on bicycle route choice and the influence of for example comfort on route choice.

#### 7.2.2 Web application

The current design is geared to a desktop application to be used by the municipality as main user. Further development of the tool towards a web application will enable the municipality to provide the bicycle information to the general public via the internet and use the information as a communication tool. Disclosing the bike information will show the municipal efforts in traffic monitoring and can help in explaining and justifying of bicycle policy measures. In appendix A the set-up of this web application is provided.

---

\(^2\) project plan i-Zone (August 7, 2015)
\(^3\) project website SUNSET (7 August, 2015)
\(^4\) project website SMART (7 August, 2015)
\(^5\) project information MMP (7 August, 2015)
7.2.3 Dynamic representation of bicycle flows
The current application is designed to be static in terms of presenting bicycle flows. The application presents estimates of bicycle flows on an average workday. A further potential improvement may be a dynamic presentation of the bicycle flows through the day, week and year. This enables studying the dynamics of bicycle flows over time.

7.2.4 Traffic predictions
The monitoring tool is currently designed as a descriptive tool presenting bicycle related information from previous years. The tool can be extended to a predictive tool for modelling future bicycle flows to assist in the ex-ante evaluation of bicycle policies. This requires the tool to allow (1) to change the network the test the addition, change or removal of a bicycle path and, (2) to alter the demographics and characteristics of the origins and destinations (e.g. introduce future housing, jobs, schools and shopping areas). Moreover, the developments in the use of bicycles should able be addressed, such as the electric bikes and the various bicycle stimulation policies. Predictions of future bicycle flows allow the municipality to evaluate the effects of bicycle policy measures on forehand.

7.3 Automation of data updates and maintenance
Currently the database is designed as fixed in terms of data and information. Updates and additions of data need to be inserted manually. The database and application does not require a real-time feed of data in the database. Currently the application is designed to be used as a long term monitoring tool. Bicycle information can be presented and compared on the basis of years. Therefore, in the design a yearly update with recent data is proposed. To help in updating the data and managing the information a tab is included in the application to load and process the traffic light data and the MON/OViN data into the relevant information. Potentially an automatic connection with the traffic light data may offer new possibilities. It may enable a more real-time overview of bicycle traffic. This requires the database to be able to retrieve data from the traffic lights and to be added to the database on a server. Currently the database is only used locally.

7.4 End-user information contribution
The database and the application are currently designed to present the bicycle information according to the requested analysis of the user. The application is currently rather one-directional and does not allow contributions of users. Adding functionality for users to comment and remark on bicycle related issues could open new possibilities for the municipality as road authority to acquire qualitative information about the bicycle traffic system. Proving the application with a functionality allowing users to pinpoint comments or remarks to a specific location in the network may enable the municipality to act faster upon bicycle related issues.
References


Divv 2010. Mobiliteit in en rond Amsterdam: een blik op de toekomst vanuit een historisch perspectief, Gemeente Amsterdam.


Morales, J., De Bij, R. 2014. Spatial database design; the model-driven architecture (MDA) approach. PowerPoint presentation. SDIT group, ITC, University of Twente.


Stadsregio Amsterdam 2015. Investeringsagenda Fiets.


Thomas, T., Jaarsma, R. & Tutert, B. 2012. Exploring temporal fluctuations of daily cycling demand on Dutch cycle paths: the influence of weather on cycling. Transportation (online version), 1-22.

U.S. Department of Transportation & Bureau of Transportation Statistics 2000. Bicycle and pedestrian data: Sources, needs, and gaps. Washington, DC

Van Ginkel, J. 2014. The value of time and comfort in bicycle appraisal. Master of Science, University of Twente, Enschede.

Appendix A: Web application design

A.1 Structure

A suitable server-side structure is an MVC design (i.e. Model, View and Controller). This is a web-based design accessible to a large audience where the information is presented on the client-side, whilst the data storage and processing is located on a server. In this way the bicycle information is shared by all users. Moreover, an MVC structure also enables the user to contribute to the information in the application. This structure could therefore be a communication tool where all users can find and contribute to bicycle related information.

In an MVC architecture the user interacts with Views, which display data held in Models and forwards the actions the user conducts in the application to the Controller. In the controller the behaviour of the application is defined. Here, the user interactions with the views are translated into actions of the application and the proper view is selected to render the results. The model represents the data used in the application and the rules for accessing and updating the data. The model notifies the view when the data changes and enables querying of the data.

A.2 Conceptual design

To support the 4 main tasks of the application an interface with 4 panels is used. A map provides the basic spatial reference. A selection panel allows filtering by relevant criteria. A results panel presents a detailed analysis of the specific selection. And the street view allows the user to validate the location of the analysis by navigating the area of interest. The conceptual design of the user interface is presented in Figure 6-3.
The map pane will be the most prominent in the web application. On the right hand side the selection panel, results panel and the street view panel are placed. The four buttons in the upper left part of the interface allow for the selection of the required analysis type. The buttons in the lower left part enable the user to customize the map panel. Figure 6-3 basically presents the views of the MVC application. In section A-3 the design of the views as well as the controllers and models of the web application are discussed in more detail.

**A.2.1 Traffic counts**

For the functionality of presenting the bicycle volumes the following requirements are formulated:

- Provide an overview of the traffic count locations in Enschede (Show links with a counter)
- Show traffic profiles of selected count location according to user-defined time period (this relates to the bicycle profiles as presented in Figure 4-2)
- Make a comparison of two time periods at one count location
- Add visual counts for a selected location

The following figure presents an example of the set-up of such functionality.
A.2.2 Traffic flows

For the functionality of presenting the bicycle flows the following requirements are formulated:

- Show origins and destinations based on the NTS data (not assigned to the network) (using either the centroids as ODs or the corresponding postal zones)
- Show traffic loads on link based on the NTS data and calibrated with the count data. Also show the count locations that were used for the calibration
- Allow for comparison between two time periods (i.e. years) based on the calibration of the OD-matrix on the average count values for the respective years
- Present a dynamic figure of the evolution of bicycle flows over a period (i.e. workday)

The following figure presents an example of the set-up of such functionality.
A.2.3 Travel behaviour of neighbourhoods
For the functionality of presenting the trip generation and modal split of neighbourhoods the following requirements are formulated:

- Show the postal zones
- Allow the selection of postal zone
- Show trip generation and modal split of selected zone
- Allow for making comparison of selected zones

A.2.4 Selected link analysis
For the functionality of conducting a selected link analysis the following requirements are formulated:

- Have a specific set of links to choose from (main bicycle tracks: intersections of radials with single, paths with generally high bicycle volumes)
- Show estimated load for a selected link for pre-defined time periods (morning peak, afternoon peak, work day, weekend day)
- Allow for making comparisons of a selected link
- Show estimated load/average count volume for a selected link (using user defined input: fromdate, todate, weekday, day type, etc.)
A.2.5 Additional functionalities
In addition to the previously mentioned functionalities the application can be expanded with the following aspects:

- OD matrix based solely on the NTS (PC4s as origins and destinations and not assigned to the network)
- Select an average daily profile of a count location in accordance with user selection
- Have pre-defined time periods (workday morning peak, afternoon peak, workday, weekend day) for the user to select from. This may require some calculations in the database
- Enable the selection of a particular travel mode (from the NTS) and estimate the associated traffic flows
- Present a dynamic image of traffic loads over a day/week/year

A.3 Final design
The web application is designed in accordance with the MVC structure. In this section the required views, controllers and models are discussed that compose the web application.

A.3.1 Views
There are four views in the application as presented in Figure 6-3 (i.e. the four general components). The map view will be the most prominent containing a map as a base layer and several overlays as containers of the spatial bicycle information (e.g. counts and estimated loads). Here public data services can be used (e.g. maps provided by PDOK\(^6\)). On the right there are three smaller views. The middle one is the selection view. Here, the user can make a selection of the information to be presented (i.e. count location or road segment, day type, year). The user can choose to construct a comparison based on all selection parameters. The upper right view will show the results of the selection, for example a chart presenting the volume profile over the day. The lower right view will

\(^6\) https://www.pdok.nl/
show the Google StreetView of the current location in the map view or the selection view as a reference to the user as to he selected the correct count location or road segment.

A.3.2 Controllers
The controller is responsible for capturing events triggered by the user (e.g. a filter selection) and retrieve the corresponding data values from the database.

The first event the controllers have to manage is the type of analysis the user wants to conduct. After the user clicks one of the four buttons in the upper left part of the interface the map panel and the selection panel need to be updated to show the appropriate information. For example, when the bicycle counts are requested, a list with count locations needs to be presented in the selection panel and the locations need to be mapped in the map panel.

The mapping controller will be responsible for interaction with external services such as Google Maps. The navigation controller will be responsible for the interaction with the map viewer. The selection controller will be responsible for the interaction with the database and the presentation of the information.

A.3.3 Models
The models in the web application reflect the structure of the data to be presented. In this application the data classes as presented in Figure 5-4. There are several types of information stored in the database. Related to the bicycle network a count location model, road segment model and postal zone model are required. The traffic data (i.e. traffic light data and NTS data) will be described in a bike profile model and bike flow model to enable the representation of the bicycle information.