Deliverable

D3.3 Smart City domains, models and interaction frameworks v1

Project Acronym:	DUET	
Project title:	Digital Urban European Twins	
Grant Agreement No.	870697	
Website:	www.digitalurbantwins.eu	
Version:	1.0	
Date:	27 July 2020	
Responsible Partner:	KUL	
Contributing Partners:	TNO, P4ALL, IMEC	
Reviewers:	Philippe Michiels (IMEC) Hans Cornelissen (TNO) Yannis Charalabidis Michiel Van Peteghem Andrew Stott Pieter Morlion	
Dissemination Level:	Public	X
	Confidential – only consortium members and European Commission	

This project has received financial support from the European Union's Horizon 2020 Programme under grant agreement no. 870697

Revision History

Revision	Date	Author	Organization	Description
0.1	14/05/2020	Paul Ortmann, Willem Himpe	KUL	Initial structure
0.2	22/06/2020	Paul Ortmann, Walter Lohmann, Chris Tampere	KUL/TNO	Draft for Review
0.3	01/07/2020	Pieter Morlion	MORE LION	Review from mobility perspective
0.4	07/14/2020	Bruno Herman, Philippe Michiels	IMEC	City Flow Model
0.5	07/15/2020	Lieven Raes, Paul Ortmann	AIV/KUL	Internal Review
0.6	22/07/2020	Daniel Beran Paul Ortmann Max Schreuder	P4ALL KUL TNO	updates based on review remarks
0.7	24/07/2020	Max Schreuder	TNO	Final version
1.0	27/07/2020	Paul Ortmann	KUL	Minor corrections, figure labelling

Table of Contents

1. Executive Summary **4**

2 Introduction **5**
- 2.1. Transport modeling in a digital twin context 5
  - 2.1.1. A range of transport modeling tools 6
  - 2.1.2. Transportation models in a digital twin context 7

3. Model Descriptions **9**
- 3.1. Traffic Models 9
  - 3.1.1. Static Traffic Assignment 9
    - 3.1.1.1 Features and operation 9
    - 3.1.1.2 Scenarios and Use in DUET 13
  - 3.1.2. Dynamic Traffic Assignment 14
    - 3.1.2.1 Features and operation 14
    - 3.1.2.2 Scenarios and Use in DUET 16
    - 3.1.2.3 Visualization Possibilities 17
  - 3.1.3. Local Traffic Cityflows 26
    - 3.1.3.1 Features and operation 26
    - 3.1.3.2 Scenarios and Use in DUET 29
- 3.2. Air quality emission model 30
  - 3.2.1. TNO model 30
    - 3.2.1.1 Introduction 30
    - 3.2.1.2 Features and operation 30
    - 3.2.1.3 Calculation method 31
    - 3.2.1.4 Scenarios and Use in DUET 41
  - 3.2.2. VITO model 41
- 3.3. Noise emission models 42
  - 3.3.1. Noise model 42
    - 3.3.1.1 Features and operation 42
    - 3.3.1.2 Scenarios and Use in DUET 42
  - 3.3.2. Urban Strategy noise model 43
    - 3.3.2.1 Introduction 43
    - 3.3.2.2 Features and operation 43
    - 3.3.2.3 Scenarios and Use in DUET 50

4. Integrating multiple models in a Digital Twin **51**
- 4.1 Low Emission Zone modelling - demonstrating some challenges of integration 52

5. Conclusion & Vision towards v2 **55**

6. Appendix **56**
- 6.1. Static traffic model (P4ALL) 56
- 6.2. Noise model (P4ALL) 57
- 6.3 Showcasing how a basic Traffic model can be created 57

1. Executive Summary

This deliverable is separated into two explicit versions with v2 to be delivered at the end of 2021.
The document at hand first and foremost gives an overview of the models that the different partners can provide to DUET. The descriptions of the models have varying levels of detail/ abstraction, part of v2 will be to make these specifications more homogeneous in form.
Meetings leading up to this deliverable have revealed that there is a mismatch between some of the epics that are being pushed and the modelling expertise of the partners. The partners involved with mobility modelling (KUL/P4ALL) use regional traffic models to predict traffic flows between different areas in a given study area, yet many of the epics that are proposed in D2.3 either concern much more local problems such as parking or involve domain models that we do not have readily available in the consortium (Public Transport/ Modal Choice). In the months going forward the modelling partners and pilots need to be involved in a conscious effort to align on feasible epics - both from a modelling - and data perspective.
The introduction of this deliverable addresses these concerns on a more conceptual level and sketches how models operate at different geographical scales. On the one hand, this is closely linked to the effort involved in data acquisition - finer grained models typically need data at a higher spatial resolution - and on the other hand it affects model validity; the functional relationships and parameters that go into a descriptive model are typically developed for a range of geographical scales and may not be valid when ‘zooming’ in or out too far. For a successful study within the Digital Twin environment the user needs to be aware of such pitfalls and choose the appropriate tools to answer his or her policy inquiries.
Chapter 4 discusses some possibilities of model interactions and addresses theoretical concerns regarding circular dependencies. We showcase a rather simple example of interaction with models available within DUET, specifically, how the short-term impact of Low Emission Zones (LEZs) could theoretically be evaluated. In v2 this section will be expanded and detail precisely which kind of interactions are supported and what the most valuable avenues for extensions are.
The deliverable also serves as a reference for the system architects that need to facilitate the data-flows sketched for the models.

2 Introduction

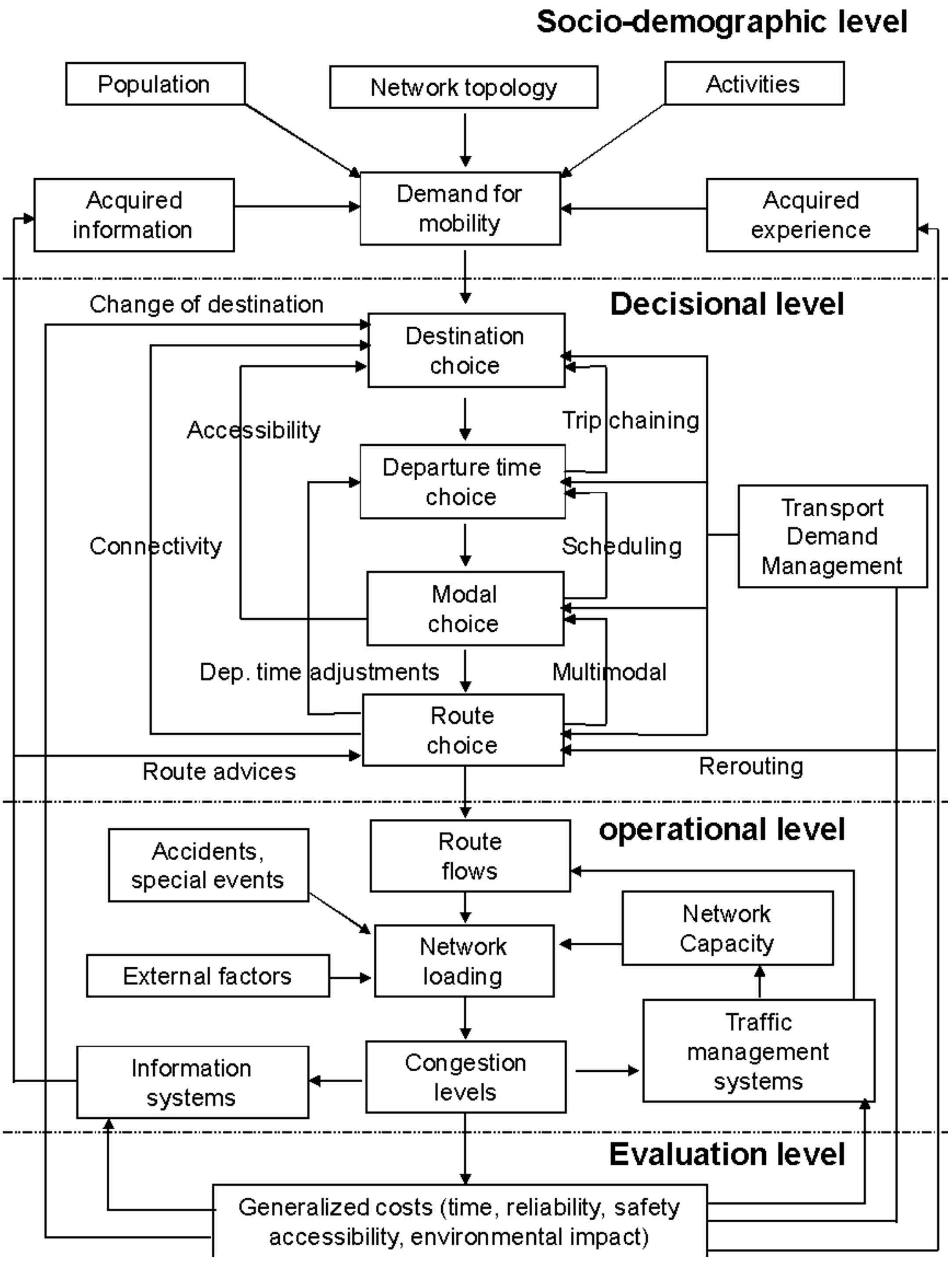
This deliverable is separated into two explicit versions with v2 to be delivered at the end of 2021. This first version describes the models that the technical partners (namely TNO, P4ALL and KUL) can provide to the digital twin (private Traffic, Air Quality and Noise models).
It serves as a reference and starting point to refine the modelling APIs that are being developed in the Alpha version. Additionally, it gives an overview of the modelling competencies of the different technical partners in the consortium. The document should facilitate a more informed discussion on the feasibility of some of the user requirements that are being posed in D2.3, both from a modelling viewpoint, e.g. if we can address the issue presented in a particular epic, and a data perspective (Are the provided data sets sufficient in detail and quality to give a valid answer for a scenario?).
The following introductory chapter gives an idea of the challenges that arise in transport modelling in the context of Digital Twins. The third chapter introduces the Traffic, Air Quality and Noise models that are provided by the technical partners. The fourth chapter discusses some possibilities for interactions between models and limitations regarding computational feasibility and validity that need to be considered. In the conclusion we state our vision towards v2 of this deliverable and address some of the shortcomings that are still present.

2.1. Transport modeling in a digital twin context

The transportation system is an open system, where travelers of different population segments interact in different transportation modes to enable different types of activities. There exists no single all-encompassing way of modeling all these interactions. Rather, a multitude of model components exist that can be combined into a whole range of transport models.

Figure 1 brings some structure into the cause and effect relationships within transportation systems, and therefore in the many existing models and model components. It shows how the demand for mobility is linked to interactions at the socio-demographic level. A whole series of decisions converts the demand for mobility into specific demands for trips along specific routes. At the operational level, finally, these trips come together in the transport infrastructures, which determines the locally observed flows and delays along every road and intersection in the network. The stakeholders involved in transportation can evaluate the resulting traffic patterns from various perspectives, like the time cost, environmental impact, safety cost etcetera. Note how many feedback mechanisms connect phenomena on one level in the transportation system to other levels, which creates dynamic interactions on the short, medium and long term. Also, it means that typically the impact of any change within the transportation system or one of its boundary conditions does not remain local, but trickles down or up to many related parts of the transportation system.

Figure 1: transportation system structure and model components



2.1.1. A range of transport modeling tools

The schematic representation of figure 1 is general. Depending on the socio-demographic level of interest, and on the evaluations that the analyst is interested in, the models providing insight into transportation can differ substantially. Table 1 lists some additional dimensions along which transportation analyses can be categorized. It is obvious that radically different transport models are required for an investigation in the short-term impact (e.g. next week) of a local deviation in a residential neighborhood, versus a 10-year impact assessment of a road charging scheme on the lower income groups in a metropolitan region. Not only will the data requirement (resolution/aggregation level) differ substantially, also the behavior affecting the evaluation criteria of interest differ significantly. This may range to coarse, aggregate regressions of how a population shifts between travel options to detailed individual decision-making models of separate travelers. Finally, the type of model can range from:

- Pure descriptive models: simple empirical regressions of observed data, over
- Explanatory models: that try to capture the relationships between stimuli and the resulting behavior, over
- Normative models: that look for optimal values or designs of certain control variables in a transportation system (e.g. optimal price for maximum socio-economic welfare, or optimal public transport design for social inclusion) to
- Predictive models: that often combine descriptive and explanatory models.

Table 1: dimensions for transport analyses

Dimension	Category	Example
space	scope	terminal, city block, city, metropolitan area, region, country, continent, world
	resolution	single building, neighborhood, municipality
time	scope	historical, real-time, short-term forecast, long-term scenario forecast
	resolution	seconds, minutes, hours, peak/off-peak, day, month, year
population	resolution	individual agents, households, socio-demographic segments (e.g. income percentiles, age groups), population
activity space	purpose	social, leisure, home-work, home-school, professional, freight

2.1.2. Transportation models in a digital twin context

Traditionally, specific instances of the wide range of potential transportation models were developed – and in many cases fine-tuned over years or even decades – by specific stakeholders in the domain. A national transportation authority would develop and maintain region- or even nation-wide explanatory models for average traffic loads during peak periods in all transportation modes and infrastructures in their jurisdiction, albeit on a rather coarse resolution of municipalities or other aggregate zones (e.g. corresponding to statistical sectors in public databases). Such models support for instance scenario and cost-benefit analyses related to large infrastructure projects or tax scenarios. Local traffic controllers would develop detailed second-to-second dynamic models of traffic operations on a signalized corridor that they are operating. They may explore what-if scenarios for better incident response (e.g. prioritize emergency services), or real-time optimization of delays for different traffic types.

In a digital twin environment, the ambition is to have a digital replica of the relevant real-world objects and aspects, here focused on the transportation system. The use cases and precise positioning along the dimensions of table 1 may be undetermined a priori and may need to be flexibly adapted and further detailed as more data comes available and more stakeholders embrace the use of the digital twin for supporting their decision making or communication of plans towards the broader public. As a consequence, no dedicated selection and configuration of modeling components can be made; the ambition should be that ideally different scopes and resolutions can be defined ad hoc for a specific use case, based on a library of available model components with different characteristics.

This is an innovation challenge that needs to be gradually developed and expanded. Not only does it require a library of modules that in principle should be compatible (e.g. finer-resolution models should be disaggregates of lower-resolution versions), it also requires different data sources and calibration procedures to be integrated in the digital twin environment.

In DUET, a first step of such integration will be explored. The idea is that some basic modeling data and functionality is available covering a large territory. The basic modeling functionality will typically be on a rather high spatial aggregation level (e.g. municipality level) and low time resolution (e.g. stationary state in peak and off-peak periods). The explanatory character is rather simple, only capturing basic transportation relations between production and attraction of activities and trips in aggregate zones. This basic model layer is explained in the Appendix. More refined models allow zooming in on city regions (diameter ~10-20km) within the larger territory. The modeling principles at this level are comparable to the basic modeling layer, but the principles have been applied on data with smaller spatial aggregates and more calibration has been performed to produce valid average (stationary) flows. This regional stationary model layer is explained in the section on Static Traffic Assignment. Traffic in regions of similar size can also be modelled in a finer time resolution in a dynamic traffic assignment model that considers peak hour dynamics with queues building up and dissolving. This model captures structural flows and routing interaction with queues, but does not provide fine details e.g. no local streets, no tracking of local queues at individual intersections along arterials, nor of local routing towards parking spaces. The model is further described in the section on Dynamic Traffic Models. Finally, the digital twin allows zooming in on local neighborhoods in city districts, where in principle every individual street and intersection is considered. Obviously, the validity and accuracy here strongly depends on the availability of detailed local data, as will be further explained in version 2 of this deliverable.

In theory, the more detailed model layers can be configured for any region or neighborhood in the larger territory. However, the usefulness of such models depends entirely on their empirical validity. At present, no automated calibration procedures exist, nor do there exist generally applicable guidelines on which data is minimally required to guarantee a certain level of validity of the model outputs. Within DUET, the more refined models will therefore be configurable in principle over the entire territory, but in practice can be trusted empirically only for those zones for which substantial calibration efforts will be performed in the context of specific pilot use cases using dedicated data sources. Conclusions will be drawn on the data and calibration requirements for other detailed models, based on the experience in DUET.

3. Model Descriptions

In this section we will introduce the different models that will be supplied in DUET to simulate traffic, noise and emissions. We will sketch some scenarios and make links to the epics that were described in Task 2.2.

3.1. Traffic Models

DUET will showcase three different kinds of traffic models, static, dynamic and a local mobility model, which can be thought of as a multi-modal traffic state estimation.

3.1.1. Static Traffic Assignment

3.1.1.1 Features and operation

Models based on Static Traffic Assignment capture two of the basic phenomena that we associate with traffic:

1. People strive to minimize their travel times by choosing the shortest route, this behaviour eventually leads to something we call ‘equilibrium’ - a state in which all the used routes from A to B have the same travel time, it’s equivalent to saying that no traveller can unilaterally reduce their travel time by switching to a different route
2. The travel time on a road increases with the amount of people that use it

We capture 1.) with what we call a **route choice model**: it specifies how travellers choose among different options to travel. We can readily imagine that route choice is more elaborate than what is described here, individuals may favor a longer route for its scenic value or to avoid slowing down for traffic lights. Route Choice Models come in many flavors, some of which are able to reproduce this variance in behavior.
2.) is represented through the **network performance model**, it describes the travel times on the roads given the travel plans of all individuals. Static models make use of Volume Delay Functions, they provide average travel times on each road given the flow over a modelling period (typically a peak). Different versions of these functions exist: BPR, conical, Akcelik, Lohse, see [1] and [2].

Static Traffic Assignment models can be trusted with calculating aggregate characteristics, but become less and less trustworthy with smaller geographical scope. They can be used to, see [3] :

- Obtain good aggregate network measures, like total motorway flows
- Estimate zone to zone travel costs (times) for a given travel demand
- To obtain link flows in the correct order of magnitude and identify heavily congested links

and less reliably to:

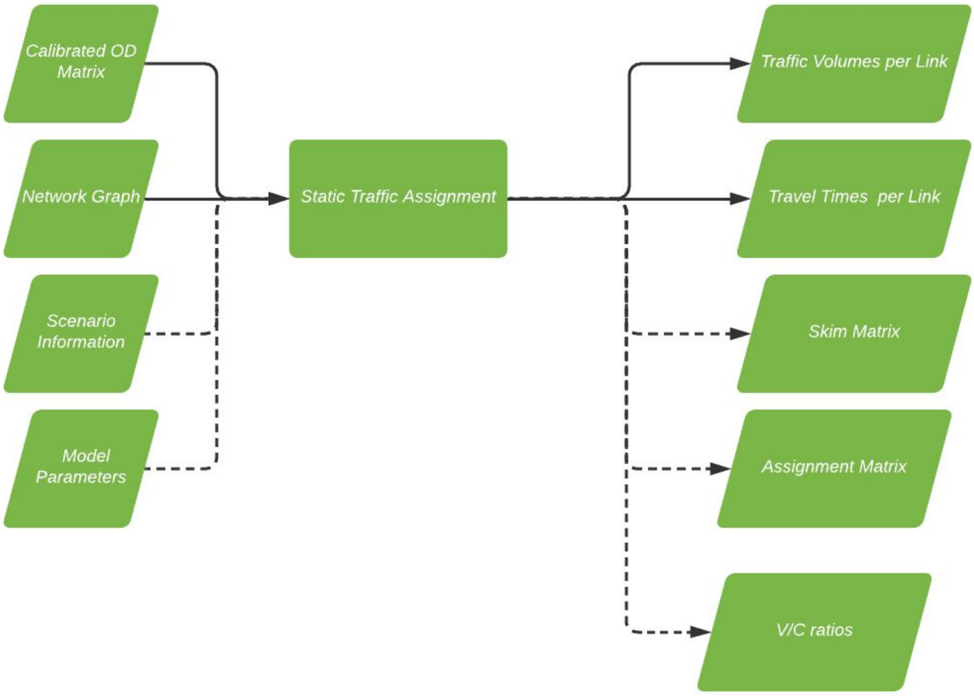
- Estimate routes used between each Origin - Destination pair
- To analyse which Origin - Destination pairs use a particular link/route
- Obtain turning movements for the design of future junctions

[1] Highway Capacity Manual, “Highway Capacity Manual,” *Washington, DC* 2 (2000).
[2] Rahmi Akçelik, “Travel Time Functions for Transport Planning Purposes: Davidson’s Function, Its Time Dependent Form and Alternative Travel Time Function,” *Australian Road Research* 21, no. 3 (1991).
[3] Juan de Dios Ortúzar and Luis G. Willumsen, *Modelling Transport* (John wiley & sons, 2011).

In a nutshell they allow us to roughly anticipate the impact of a change in either network characteristics or/and travel demand. How exactly that change is to be evaluated depends on the purpose of the study. A often cited generic goal is the reduction of total time spent by all travellers.

Inputs and Outputs

Figure 2: Data Flow Static Traffic Assignment



broken lines indicate optional I/O

Network Graph

Traffic models require a directed graph structure to operate on, edges represent streets/paths that different vehicles have access to. Each included edge in the graph should represent a road that can be used by cars. For each road we need to know free-flow speed, capacity, toll (if applicable) and direction (one way or not).

Hierarchical road network graphs based on OpenStreetMap, a Volunteered Geographic Information System, can be generated with open source software such as OSMnx, a python package. It's noteworthy that the quality of the datasets that are generated in this way vary widely between regions and depend on the diligence of the local contributors, see [4].

Capacity in traffic modelling is defined as the maximum flow that we can observe on a given road over time (in vehicle units/hour). It's an input into the previously mentioned Volume Delay functions, typically these apply to links, but turn and node impedances can also be incorporated here - the approaches differ between signalized and non signalized intersections. Currently P4All allows passing these turn delays as fixed values.

In Static models capacity is often a parameter that is integrated into calibration efforts and helps to describe the delay incurred at the intersections that are at the end of a link.

[4] Jean-François Girres and Guillaume Touya, "Quality Assessment of the French OpenStreetMap Dataset," *Transactions in GIS* 14, no. 4 (2010): 435–459.

The impedance functions for different road classes may differ, it's common in traffic models to assign different 'link types' based on this hierarchy and calibrate the parameters of the link impedance function for each class.

A node indicates either a change in edge characteristics - different capacity/ speed - or/ and an intersection if more than two edges are connected to the same node.

For each node we need a sort of binary connection matrix indicating feasible turns, it simply specifies if it's possible to turn towards road *j* coming from road *i* for all possible combinations at a given intersection.

Calibrated Origin - Destination Matrix

Typically Travel demand is described in aggregate terms, parts of a city/ region are classified into zones and travel demand patterns are described as flows between those zones. Zones can be any geographical shape from residential districts to large structures like university buildings. What they have in common is that they either attract or produce the same order of magnitude in traffic.

An origin - destination matrix (OD matrix) gives flows for any pair of zones, in Static Traffic Assignment these flows are typically indicative of the demand over the peak period. This is often downsampled to represent the flow of one hour during that peak.

The starting point for creating these matrices is a trip generation model that determines the production and attraction of these zones; a more detailed description can be found in the Appendix.

Creating these matrices from scratch is no easy task and requires combining this geographic data with socioeconomic data (such as car ownership rate, income ..) or/and travel surveys. Since the availability and nature of the data that is available for this varies widely between regions/cities there is no straight-forward way of automating OD matrix generation.

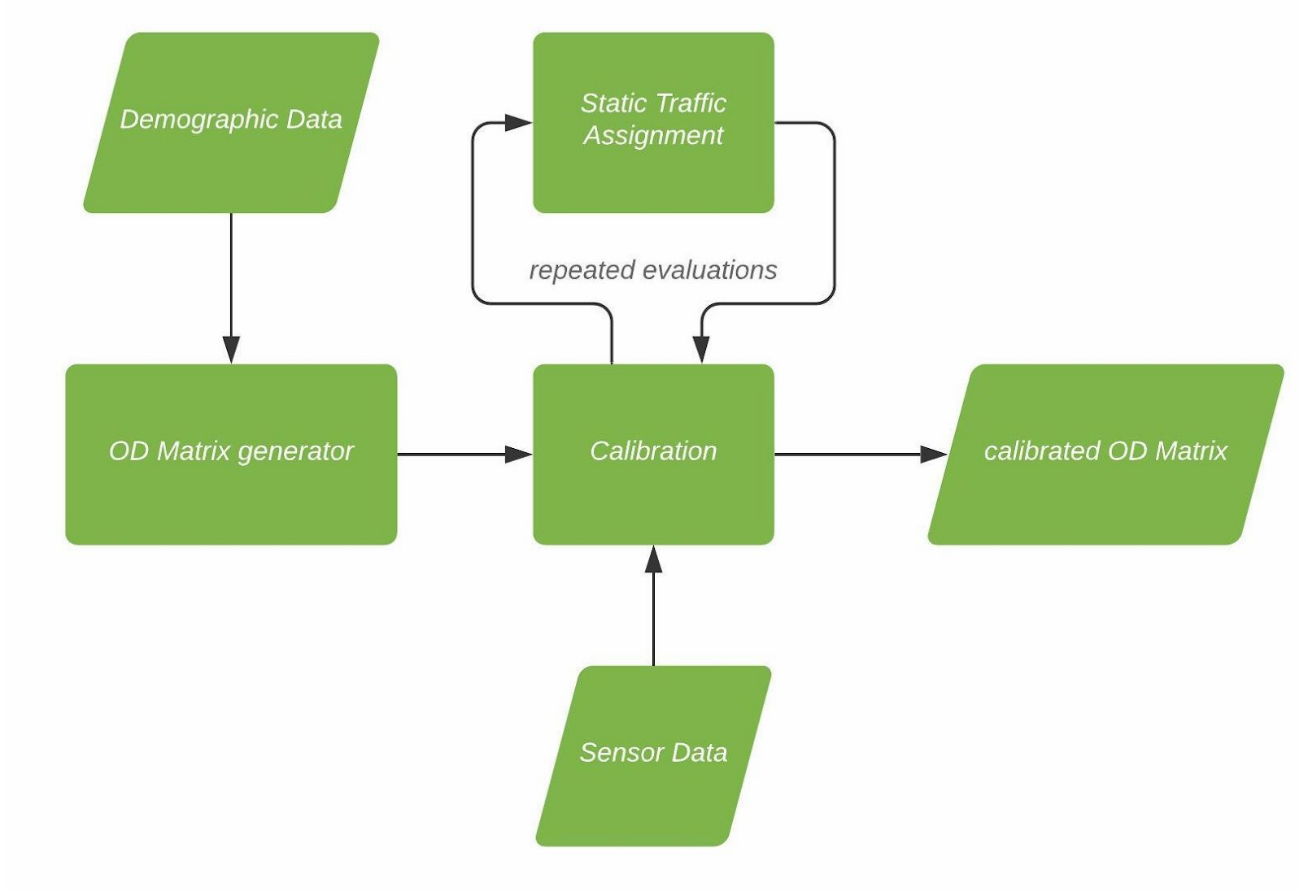
In figure 3 below you see this first step and a rough scheme of how (demand) calibration works. The results of a scenario analysis using Static Traffic Assignment can only be valid if the mobility flows in the system somewhat mirror reality, therefore it is essential to correct the initially created OD matrix by using sensor and/or traffic detector data.

A calibration procedure typically repeatedly evaluates a Static Traffic Assignment given an altered OD matrix, until the flows that we observe on the links with sensors/detectors are close to what is predicted by the assignment procedure. This explanation here simplifies things drastically, some of the issues that may arise are: Sensor data may not represent an average day (accidents, construction, storm), (local) density of the sensors not sufficient, ⋯

Similar calibration procedures exist for adjusting the average travel times that Static Traffic Assignment gives for each road. These travel times may be obtained by matching vehicles that cross different sensors individually, or through individual GPS data. This can then be used to calibrate the parameters in the impedance functions for different link types.

Calibration is a computationally expensive task, may take several days and will typically have to be done offline.

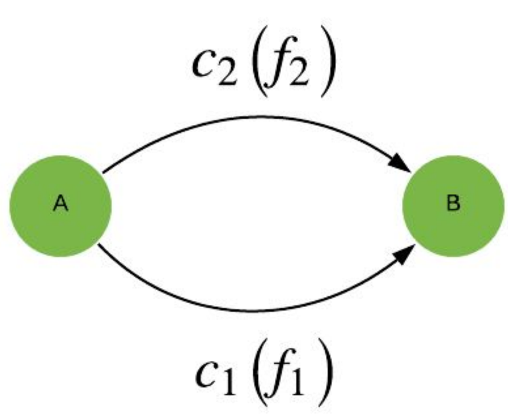
Figure 3: Demand Calibration in Static Traffic Assignment



Model Parameters

Static Traffic Assignment is an iterative procedure that (ideally) reaches the point where the cost difference between any used path is zero (see opening of this section). In reality the underlying algorithms will typically not reach this point in a reasonable amount of time. Convergence criteria are used to abort the procedure when the solution is good enough for the intended purpose.

Figure 4: Cost dependence on flow



This is an optional input for advanced users: the gap. Usually the developers of the model will have provided a default value for this, but it can be overridden by the user.

We illustrate the concept of convergence on a small example and then motivate the need for a gap measure for the entire network. Take a look at figure 3: c_1 and c_2 indicate the cost that is experienced on the two routes connecting A and B. It's a function of the flow over each of the routes, denoted here with f_1 and f_2. Let's assume that initially one of the routes is faster than the other with no travellers on either route e.g. $c_2(f_2 = 0) > c_1(f_1 = 0)$. Then it's clear that travellers going from A to B would take route 1 as it has a lower travel time (not assuming stochasticity in choice). However, since the costs on both routes are functions of the flow going over them we can find a demand

level D at which route 2 becomes equally attractive[5], that is: $c_2(f_2 = 0) = c_1(f_1 = D)$. For any demand exceeding D travellers will use both routes instead of only one of them because they aim to minimize their travel times. A gap of zero means that we've found the perfect split between the two routes that equalizes cost. (no traveller can unilaterally switch routes and reduce their travel times)
The problem shown here can be solved analytically, but real networks are more complicated. In fact both routes will be a series of links in the road network graph each with individual cost functions and the travellers using those roads have differing travel plans.
You may take a look at one of the subproblems: loading travellers on their routes and equalizing the cost for one pair of origin and destination as depicted. However if you then solve the next subproblem for another such pair (say C and D) your solution may have disequilibrated the previously computed one for AB! If any of the roads that you used for your solution to the AB subproblem[6] are also used in the CD subproblem the cost between AB will most likely no longer be equal among used routes. Algorithms solve these problems repeatedly and approach equilibrium after many iterations but these dependencies remain. The gap is an indicator of the size of the cost differences that still exist between used paths in the model results for the entire network. Different definitions exist (relative, average excess cost, maximum excess cost), which is why we keep this abstract here.
Other Model Parameters that are typically provided in calibration runs are the parameters of the impedance functions for the different link types (if applicable).

Outputs
As shown in Figure 1 Static Traffic Assignment yields flow and average delays over the provided time period for each road. The average delays for all OD pairs are summarized in what is called a skim matrix. Volume over capacity ratios (V/C) are used to give an indication of the level of congestion on the links. They're often translated to a color map to make it easier to identify congested sections in a city/region.
In post processing additional data about which origin - destination pairs contribute to the flow on each link can be calculated. This is captured in the Assignment Matrix, it is often used for calibration and to get more insights into the flows crossing particular links (flow bundle analysis, selected link analysis, see visualization examples).

Reference days
The data utilized in calibration may come from numerous sources, it is essential to make sure that it has all been collected on the same day and under the same conditions. External factors such as accidents or bad weather do affect traffic and the capacities that our transportation system has drastically. If we build up models based on data that was gathered on atypical or different days the predictions that our model can make become less trustworthy.
Road authorities address this by building up databases and observing traffic over a longer period. Data gathered on days without any incidents or atypical bottleneck formation serve as input for building models for a reference day or reference week.
If you want to predict the traffic state in a city for a festival that takes place on the weekend your baseline scenario (a weekend day) is radically different from any week day.
More technical details of P4ALL's implementation of Static Traffic Assignment is provided in the Appendix.

3.1.1.2 Scenarios and Use in DUET

[5] assuming monotonously increasing travel time functions, e.g. the more cars use a road the more congested it gets and the more time it takes to traverse it
[6] There are different ways of partitioning the Traffic Assignment problem, what is described here is a path based approach.

With properly calibrated demand data Static Traffic Assignment can provide a useful overview over the traffic flows in the system and facilitate predictions given alternate road network graphs or different demand patterns. This covers a wide range of scenarios that should be supported by the Digital Twin. We give a non exhaustive list of possible scenarios here:

- Mobility effects of new developments: Building a new Apartment complex alters the mobility demand and causes new traffic from this location to typical destinations (work zones, universities ..)
- Tolls inducing deviations from routes on which they are levied[7]
- Simulating the impact of road blockages/ closed lanes due to construction work
- Analyzing the effects of changes to a signal plans
- Predict the changes in travel time induced by larger travel demand to a city because of an event (convention, festival ..)

Note that, to handle some of these scenarios validly extra development work, data and documentation will be needed. We cannot expect to answer all of these questions with the same ease of use and confidence as existing commercial traffic modelling software.

The outputs of Static or Dynamic Traffic Assignment Models are needed as input for various other models:

- In conjunction with a modal choice and public transport assignment model they can be used to predict changes in modal split/ travel times induced by alternative schedules/ line plans/ public transport networks
- A parking model would need the local transport demand (a 'cut-out' of a traffic model) as an input and may interact with a traffic model as a controller (parking route guidance)
- Air Pollution models
- Noise Models
- Etc..

A (private) Traffic Model is a crucial building block of a digital twin but will need to be complemented by other domain models to allow answering some of the most important issues that cities face in this day and age.

3.1.2. Dynamic Traffic Assignment

3.1.2.1 Features and operation

Dynamic Traffic Assignment (DTA) brings much more nuance by adding a temporal dimension to Traffic Assignment. It enables us to capture more aspects of the traffic system's behavior that we observe in reality. Queues are modelled explicitly and spillback effects are taken into account.

The I/O scheme depicted in figure 2 remains essentially identical but every in- and output will now have a (discrete) temporal dimension. This allows for more refined and realistic representations of the expected traffic state (see visualizations at the end of this section) but also requires more detailed data regarding intersections and typical bottlenecks.

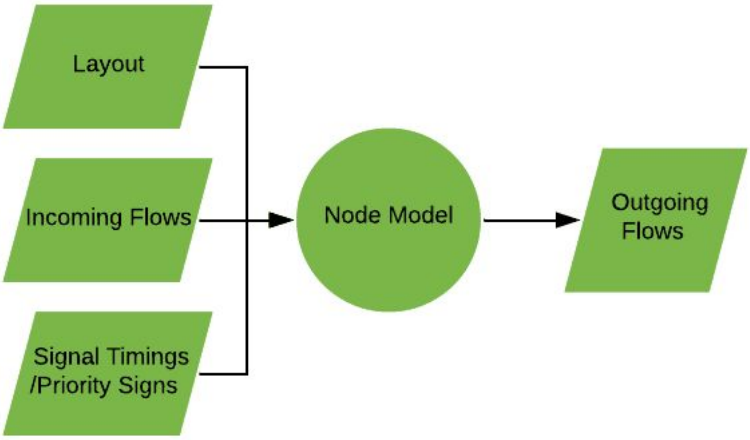
Network Graph Information

DTA requires the same kind of directed-graph structure as Static Assignment, however attributes may now vary over time. Additionally, we need fine-grained data on all the intersections in the network. Signal timings, intersection layouts, specific turn lanes and priority signs (if applicable) are needed to accurately

[7] The results of this may be overestimate traffic flows in static models. From field experiments (Jonas Eliasson, "The Stockholm Congestion Charges: An Overview," *Stockholm: Centre for Transport Studies CTS Working Paper* 7 (2014): 42.) we know that demand may respond elastically to cost increases, e.g. some people may just opt to not travel. At the same time we have no modal choice model integrated into the traffic models presented here. That would be needed to predict the switching to other modes such as public transport accurately. Nevertheless, if elasticity- and modal choice models were provided to the digital twin they could simply interface with the presented static and dynamic models.

predict the flows that cross an intersection over time, and the delays resulting from their interactions at the intersection.

Figure 5: Node models I/O



Lane changing behavior in merging, diverging or weaving sections restricts throughput and may be the cause of congestion, a bottleneck. The capacities of these bottlenecks are difficult to estimate because they are the result of the combined microscopic behaviour of all travellers that cross that particular weaving section. They are however important, as they typically cause recurrent bottlenecks creating long and long-lasting queues that trigger rerouting that may affect large parts of the region under study. In order to maintain a valid DTA proper calibration of these capacities is crucial.

Calibrated Origin - Destination Matrices

Multiple OD matrices, one for each time slice are required. The amount depends on the discretization (typical values for our model: 5 - 15 minutes per time slice). Some DTA models incorporate departure time choice and adjust these matrices as part of their calibration. This functionality is not available yet in the DTA model deployed in this project (although independently from DUET, such addition is being developed and, once ready and validated, is easy to integrate through the same API's).

Outputs

The outputs are the same as for Static Traffic Assignment with the extension that we have these link flows and impedances for each individual time slice.

The key differences between Static and Dynamic models are:

Table 2: Differences between Static and Dynamic Traffic Assignment

Static	Dynamic
Typically used to model an individual peak period, outputs like travel times represent averages over that peak	The travel times on the roads vary with time, people make decisions based on their experienced travel times
In Static Assignment vehicles occupy all roads in their route at the same time (steady state flow).	Results are discretized in time, we can estimate how many vehicles are on each road at any moment in time and predict congestion patterns

coarse approximation of delays induced by queueing	Dissolution and settling of a queue are modelled as actual processes based on traffic flow theory (depends on the network loading model, in our case it's the Link Transmission Model[8])

There are many different variants of DTA, some of which facilitate analysis that others may not.
Our DTA keeps track of the destinations of individual vehicles that travel on a link, e.g. we can provide destination distributions for every link in each time interval.
Some more system relevant effects that Dynamic models can capture and Static models do not:

- Spillback of queues onto upstream intersections and links
- Interactions between traffic on lanes (depending on the model)

This comparison is not exhaustive, it is just meant as a starting point to give an idea of the key differences and lay the groundwork for the following specification. For more detail and background on Dynamic Traffic Assignment, especially for those who already have a background in Static Assignment, we refer to.[9]

3.1.2.2 Scenarios and Use in DUET

All the scenarios that we're posed for Static Assignment are obviously also supported here. However, with a properly calibrated dynamic model the results will be more trustworthy and insightful - queue buildup and dissolution can be observed as shown in the visualization examples in the Appendix.

There's no real bound on the amount of scenarios that could be conceived with DTA. Yet, there are some scenarios that are much easier to support than others.

Any scenarios that require passing on a different/ altered network graph are straightforward and simple. This could include changing access lanes for a time period that was elaborated before the scenario, closures of roads for construction work, different speed limits, new infrastructure projects, a new traffic signal plan,⋯

Passing on a different demand pattern is also possible, much harder however is to construct demand scenarios that are valid. Calibration of dynamic OD matrices in DTA is an open research question. It's difficult to ensure that a given demand scenario is valid.

Traffic Management applications such as Ramp Metering, Route Guidance or signal coordination optimization need algorithms of their own to find the optimal solutions and may only utilize parts of DTA as a way to evaluate a solution. Those optimization mechanisms are beyond the scope of DUET, but can play a role in future, more detailed models used in a Digital Twin.

The data requirements for DTA are high, it seems unlikely that DTA can be deployed for more than some select locations or, if so, only with crude estimates of some crucial inputs (no valid results). The data set inventory does not contain some of the inputs that are needed (time dependent ODs, intersection data, capacities of bottlenecks..).

The effort involved in calibration and data acquisition and the expertise needed in model setup is significant and makes it unlikely that cities would be able to exploit Dynamic Traffic Assignment without close collaboration with traffic modelling experts. Typically, model developers and cities work together closely to obtain a valid model and successful project, see[10].

Vision for research on parallelized traffic assignment algorithms within DUET

[8] Isaak Yperman, "The Link Transmission Model for dynamic network loading," June 2007, https://lirias.kuleuven.be/1749225.
[9] Yi-Chang Chiu et al., "Dynamic Traffic Assignment: A Primer," *Dynamic Traffic Assignment: A Primer*, 2011.
[10] Yi-Chang Chiu et al., "Dynamic Traffic Assignment: A Primer," *Dynamic Traffic Assignment: A Primer*, 2011.

The city of Antwerp already has implemented a DTA model within PTV Visum (a commercial traffic modelling software). A single run of that model (typically what you would do for a what - if analysis) currently takes 1 full day, Calibration of the same model takes 100 runs e.g. around 100 days. One of our ambitions in the scope of DUET is to experiment with clusters to speed up computations in Dynamic Traffic Assignment.

This option has already been explored theoretically in the literature, see [11]. Our ambition is to extend this work in two ways:

- finding efficient (network) partitioning schemes for the traffic assignment problem. This will not only allow for faster computations of city models but allow for DTA to be deployed in larger cities, something that is currently not possible due to complexity
- explore how traffic models on larger regions can be made more computationally tractable through reductions on shortest path calculations

The second point may seem a bit abstract: If you were to construct a traffic model for Belgium and want to compute the shortest path between two points that are on opposite sides of the country the number of paths you have to consider is very large as any path crossing through a small city is a viable option to compete with the highway routes - yet we know that most travellers do opt for either highways or other roads that are higher up in the road hierarchy. We want to exploit this by reducing the search space for our shortest path algorithms to major roads for some parts of the journey, but probably not the access to the major roads from the origin and likewise the route from the highway exit to the destination.

HPC use for traffic models

KU Leuven has access to the Flemish HPC, however we are only one of many institutions that is allowed to use this facility. Access rights are required to run computations and each user has a limited budget of computing time for each month. Computational jobs are executed approximately[12] in the order in which they arrive in a centralized queue.

These restrictions, together with the long calculation time of the described models necessitate us to have most of our heavy computations offline and separate from the DUET interface. This means that HPC will be used to calculate a reference assignment with typical demand for a larger region. The results of this assignment are stored within the DUET framework and the relevant parts of it are retrieved when a model run with different parameters is requested from within DUET. We then do warm-started recomputations using the results of the assignment from the HPC within DUET.

The techniques used to speed up model calculations with a limited set of changed input parameters by use of previously computed equilibria are explained in [13].

3.1.2.3 Visualization Possibilities

Below you can find some of the typical visualizations that are generated based on (dynamic) Traffic Assignment models. These were either generated through our traffic modelling toolkit in matlab (Open Traffic Center) or within PTV Visum.

For DUET we will need to find a sufficient subset of these visualization options that are valuable to the end user and support the final epics.

Scope:

- Typically limited to the Peak period + loading/unloading of the network (5h-12h or 14h-22h)

[11] Willem Himpe, Romain Ginestou, and MJ Chris Tampère, "High Performance Computing Applied to Dynamic Traffic Assignment," *Procedia Computer Science* 151 (2019): 409–416.

[12] For more details on job scheduling on the Flemish Supercomputer see here

[13] Willem Himpe, Ruben Corthout, and M. J. Chris Tampère, "An Efficient Iterative Link Transmission Model," *Transportation Research Part B: Methodological*, Within-day Dynamics in Transportation Networks, 92 (October 1, 2016): 170–90, https://doi.org/10.1016/j.trb.2015.12.013.

- Highways, major roads, roads with important traffic function
- Modes: car, light truck, heavy truck -> PCU: person car unit to aggregate them

Visualizations

- time dependent e.g. 6:00, 6:15,⋯, 8:00,8.15,⋯12:00
- Aggregation of full analysis period
- Link level (and Node level) vs route level vs network level

Link level

Figure 6: Volume in veh/hour on each link

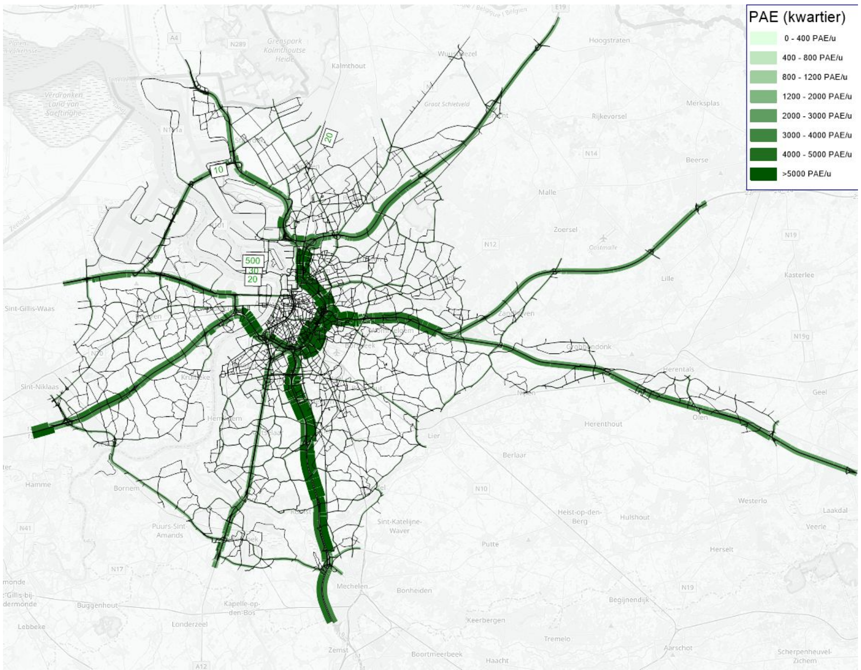


Figure 7: Density (PCU/km)

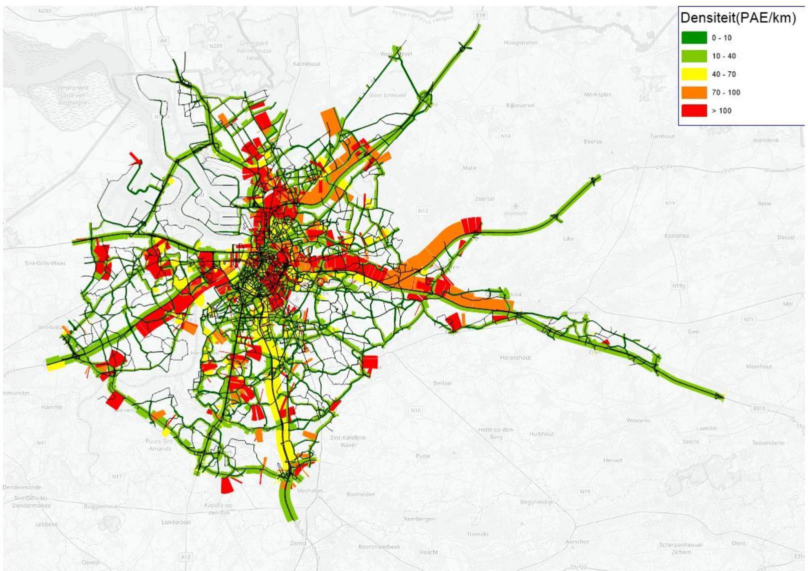


Figure 8: Relative Density (% of critical density)

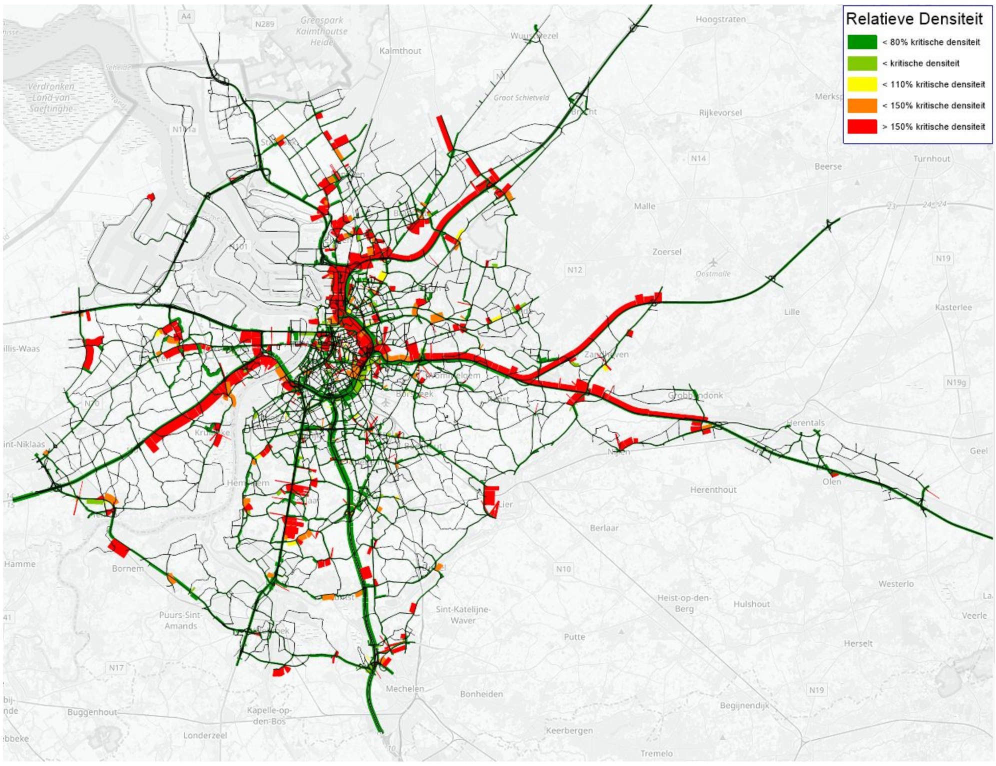


Figure 9: Speeds (km/h)

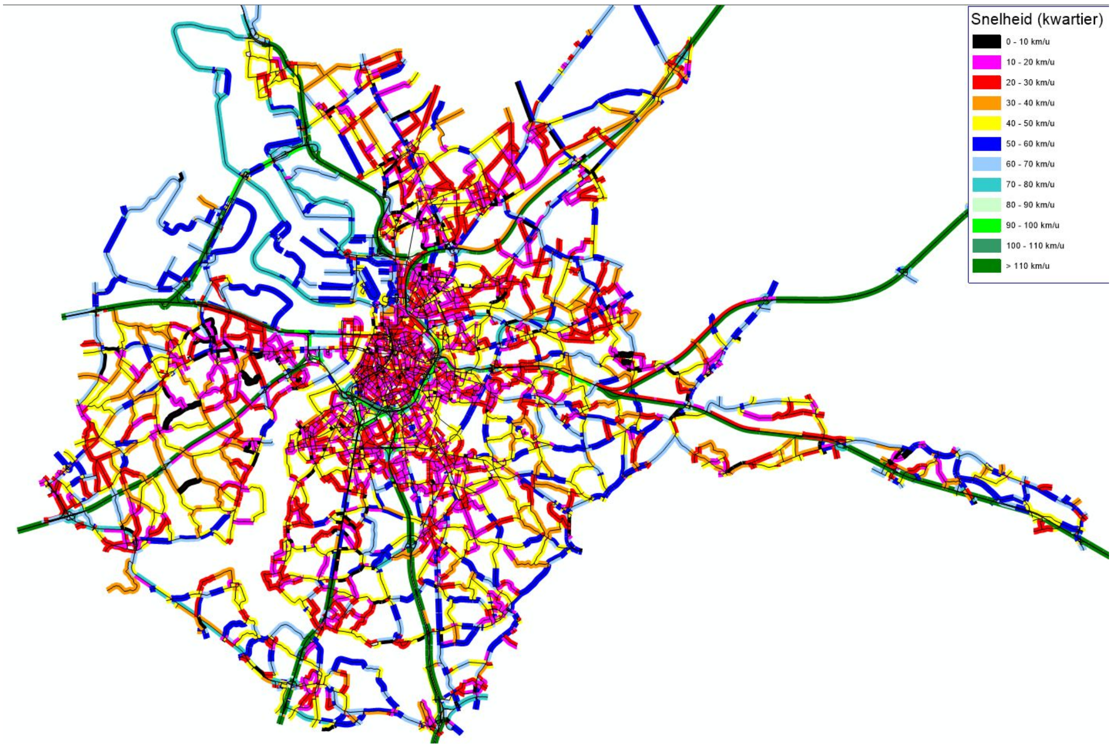


Figure 10: Degree of delay (% of free flow speed)

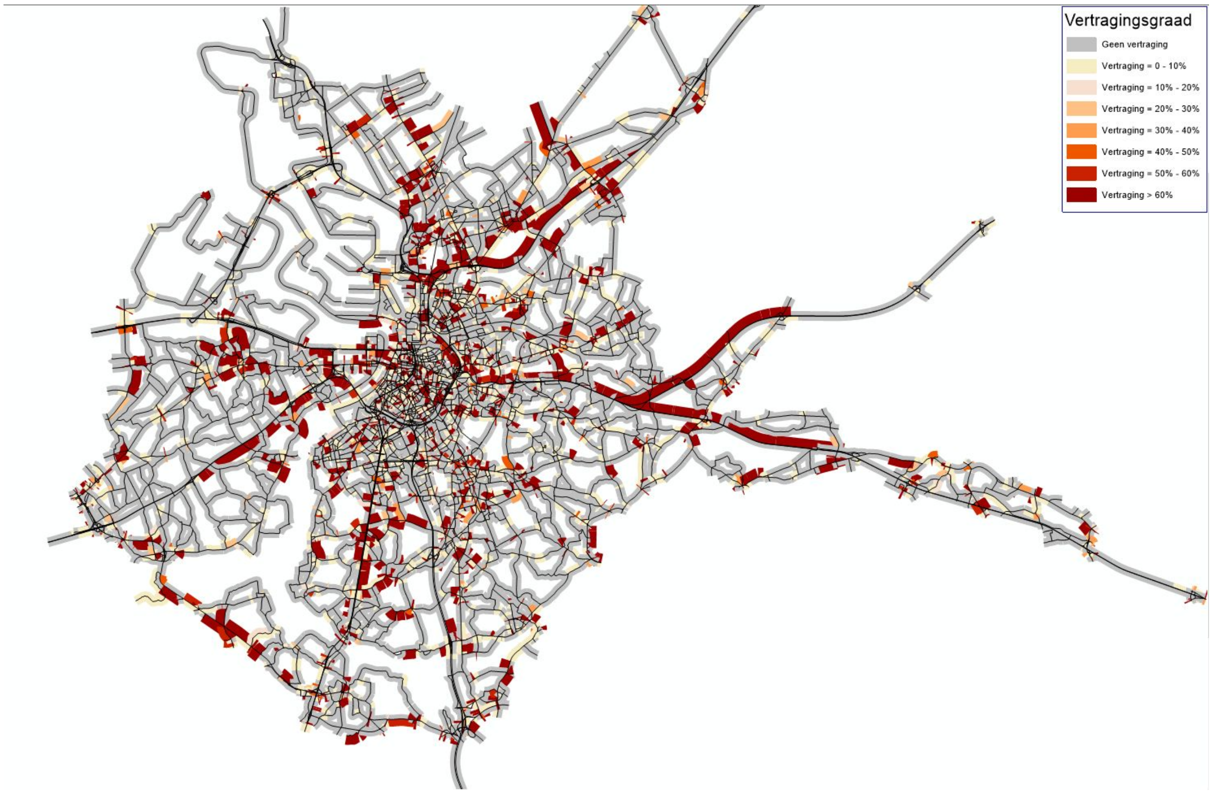


Figure 11: Detailed analysis based on link selection

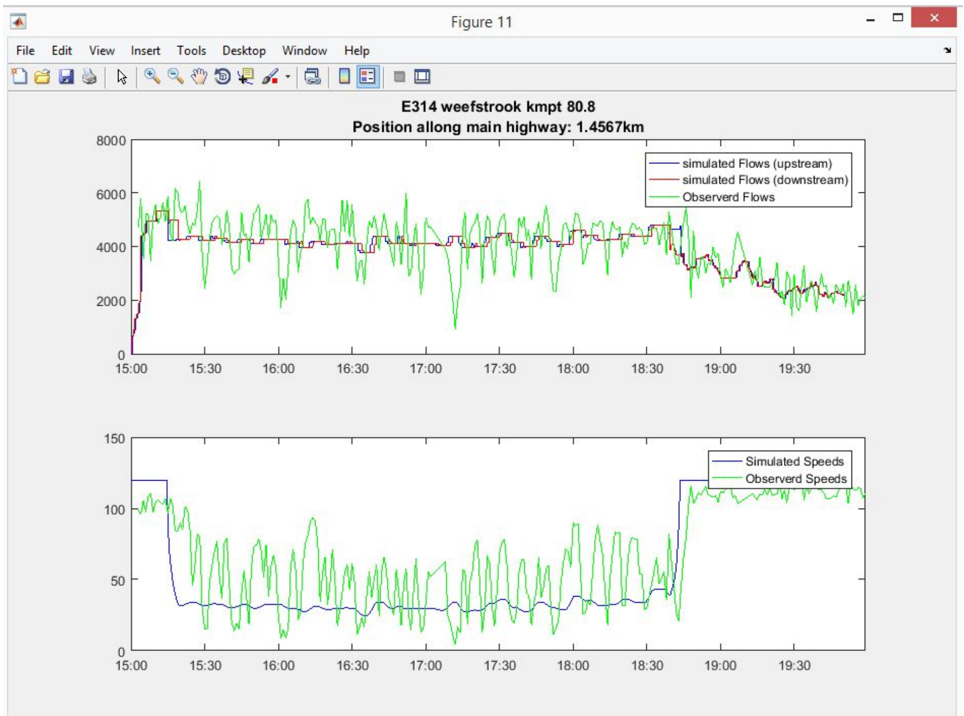
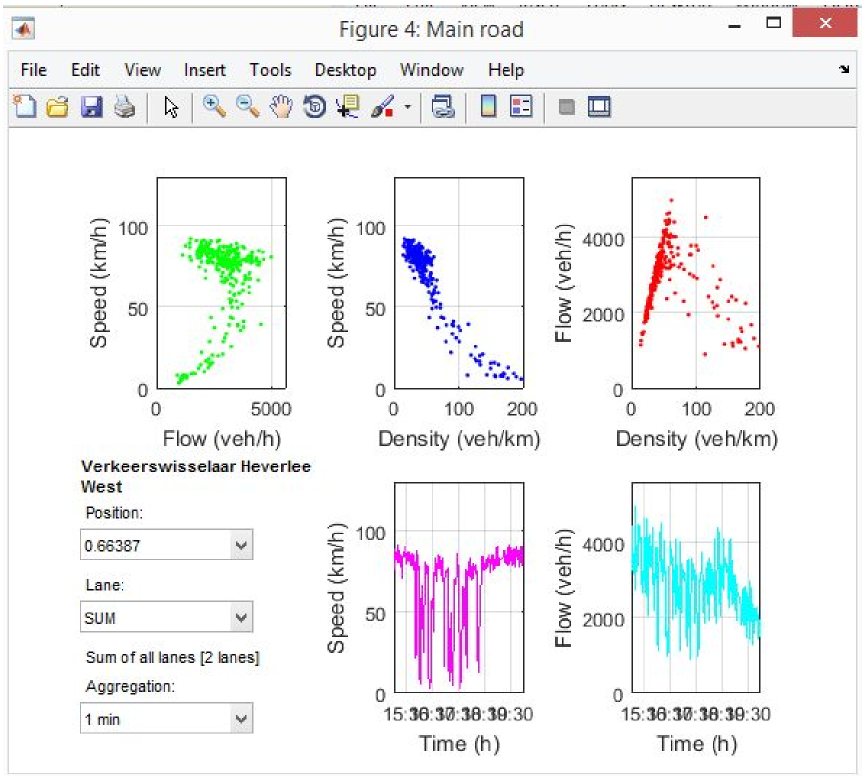


Figure 12: Fundamental Diagram (relation between flow/speed/density)



Node level

Figure 13: Volume per turn direction

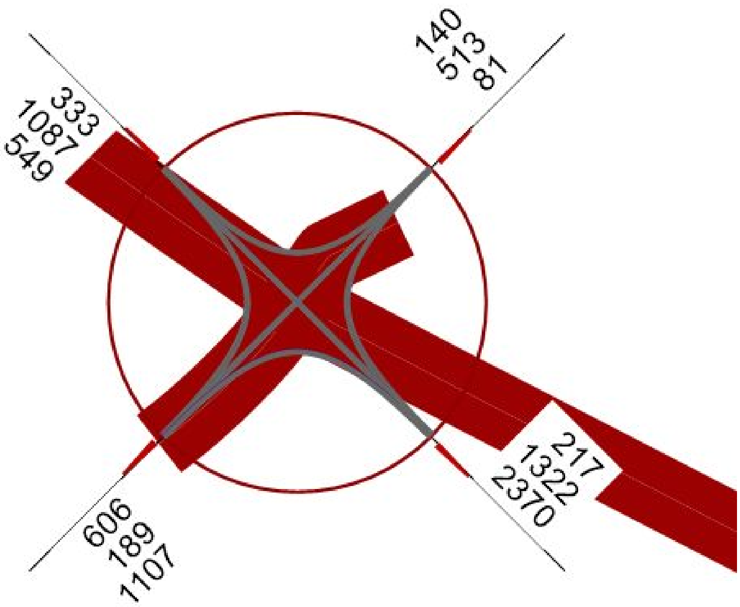
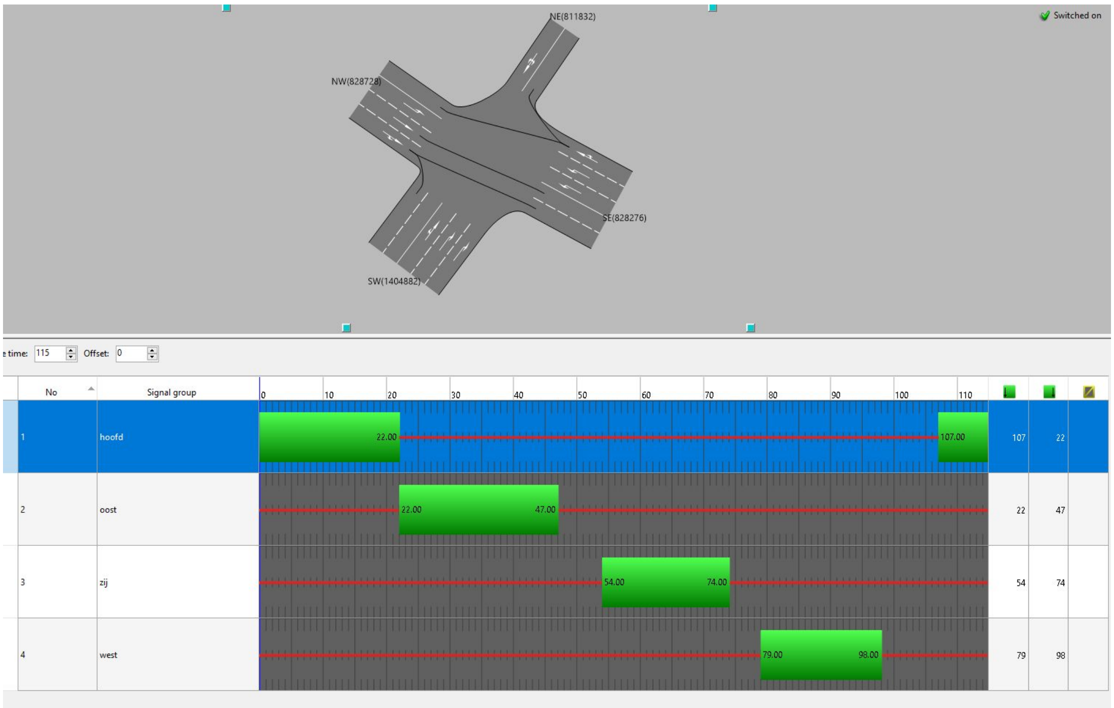


Figure 14: Intersection layout + signal timings + allowed turn movements per signal group



Route level

Figure 15: Travel time (minutes)

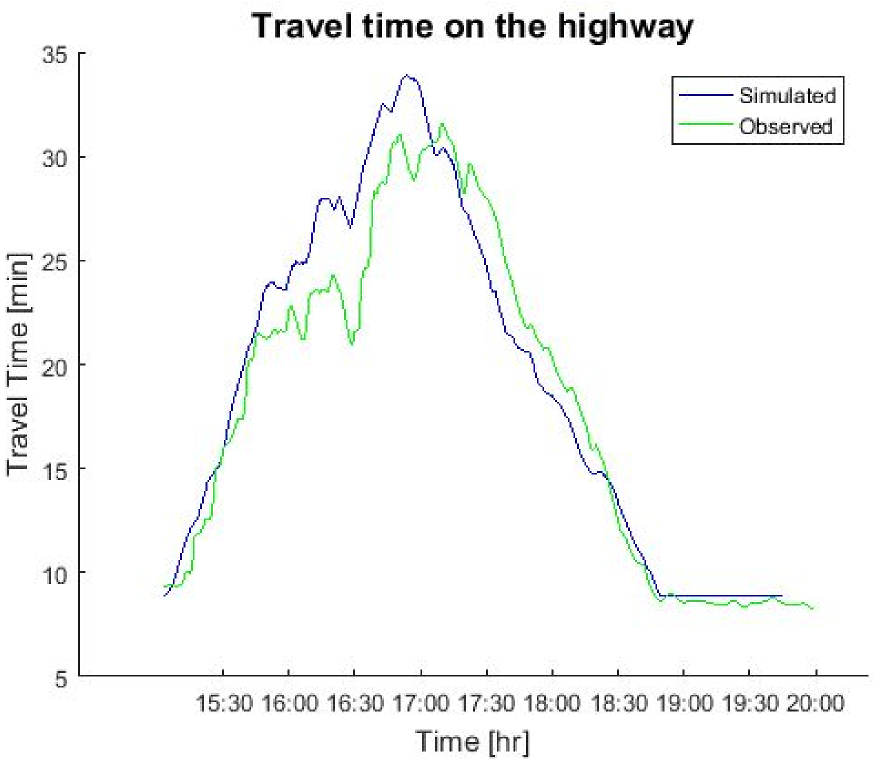
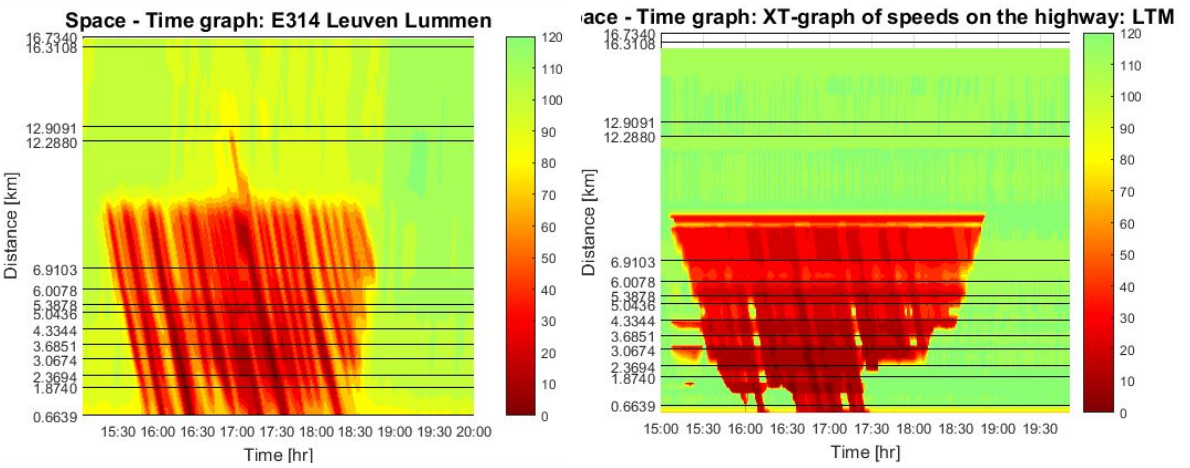


Figure 16: XT-plot (speeds km/h)

Observed Simulated



Network level

Figure 17: Desire lines (PCU/h between zones of interest)

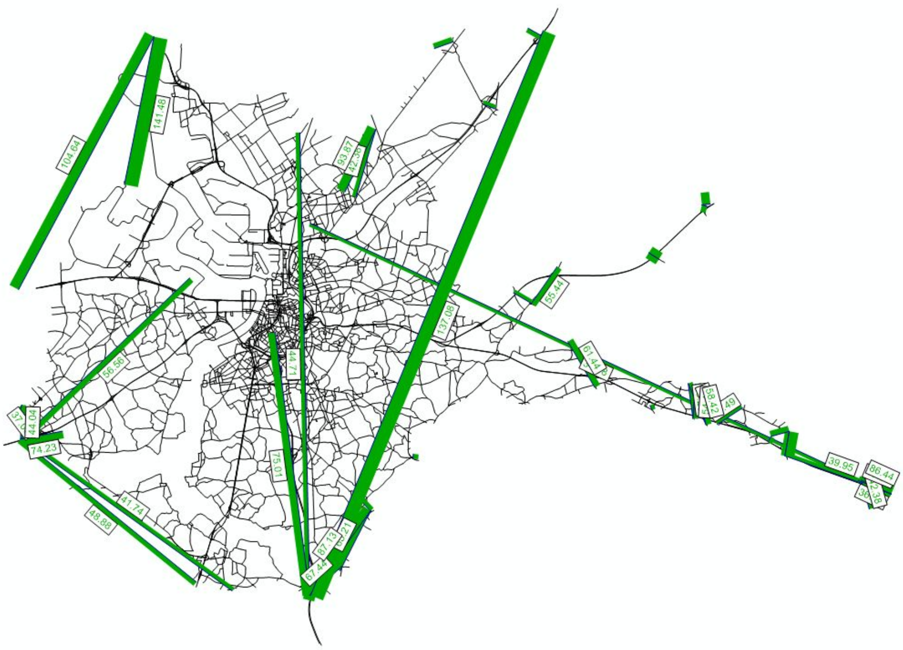


Figure 18: Accessibility (travel time - minutes)

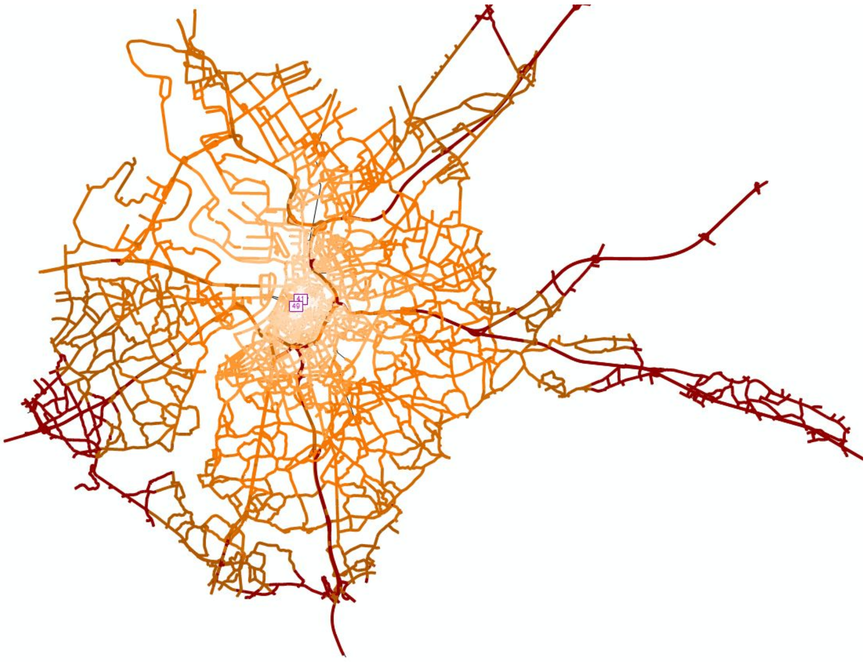


Figure 19: Selected link analysis (PCU/h for all traffic going over a selected link)

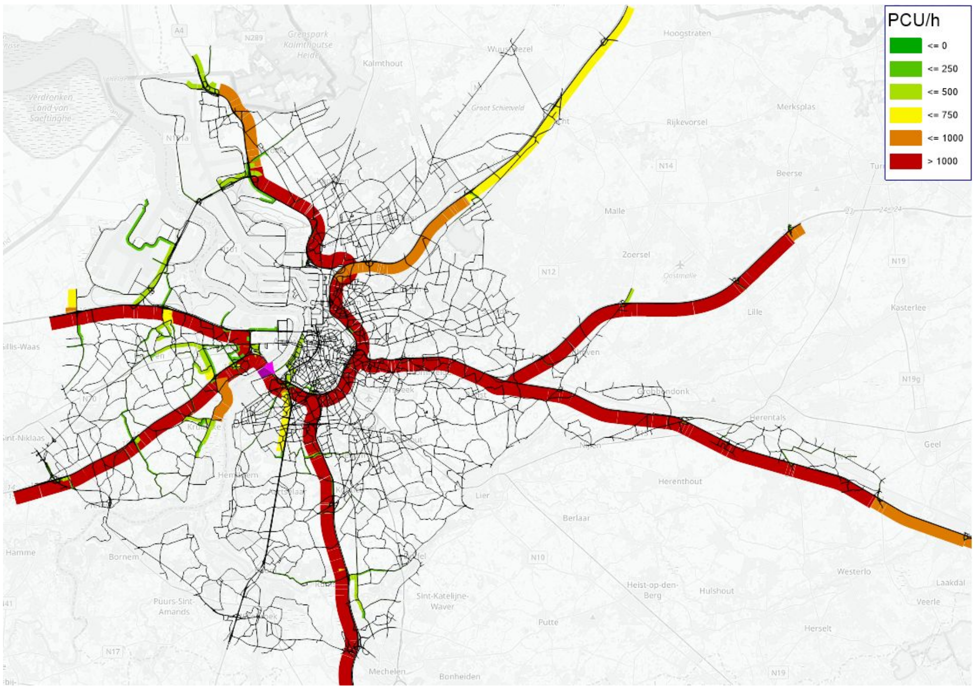


Figure 20: Multiple links selection (PCU/h for all traffic going over a set of selected links or avoiding a set of selected links)

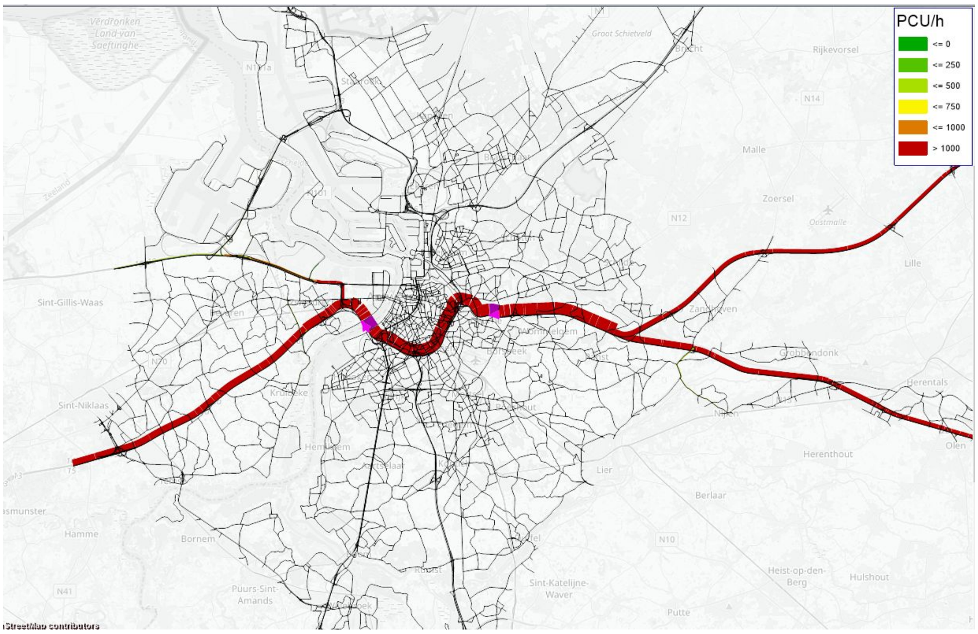


Figure 21: Scenario differences (Total volume analysis period)

3.1.3. Local Traffic Cityflows

3.1.3.1 Features and operation

Cityflows aims at using different real-time data sources in the city to better understand multi-modal dynamics of city traffic. Based on different data sources such as telco signalling data, wifi scanning data, camera object detection, Telraam data, ... the model estimates the density of traffic and discriminates between different modalities, such as motorized and non-motorized traffic, in the streets of a certain area.

Towards v2 of this deliverable this section may be expanded by literature study and framed within existing work on traffic/ mobility state estimation; see [14] for an overview of different methods on highway state estimation and [15] for an example of how traffic propagation models can be used in conjunction with Kalman Filters to obtain more accurate urban traffic state estimations based on a limited set of sensor measurements.

Inputs

The CityFlows model requires data sources to feed to its model. These can be divided into 2 classes:

- **City infrastructure:** The data on the infrastructure of the city should - as a minimum - provide a street grid in a graph-like structure, preferably in WGS84 format. Street segments and their connecting intersections are essential. Statistics about demography, population density, can all be useful information to enrich the model.

- M**obility and crowd counts:** This includes any data that gives a sense of the amount of people and/or moving direction in a certain area. A typical data source contains the following key attributes:
    - Type of data source: according to the way data is exposed, counts can be snapshot-like, cumulative with unique counts, cumulative with non-unique counts, or point-measurements.
    - Area covered by the source or sensor: For cameras: this can be the viewing angle and reach, for telecom providers it can be the Voronoi grid of a triangulated signal, for a telraam it can be the street segment covered by it, etc.
    - Time interval: especially for cumulative counts, the interval of the aggregation should be available in the metadata.
    - Modality information (optional): if data specifically measures one or more modalities: cars, bikes, pedestrians.
    - Counts: the actual payload of the data.
    - Direction (optional): a direction (if known) of counts in a street can be given.

An example NGSI-v2 TrafficFlowObserved message that we use can be seen below:


```
{
   "id": "urn:ngsi-v2:cot-imec-be:trafficflowobserved:proximus-100mE39297N31371",
   "type": "TrafficFlowObserved",
   "refDevice": {
     "type": "Relationship",
```


[14] Toru Seo et al., “Traffic State Estimation on Highway: A Comprehensive Survey,” *Annual Reviews in Control* 43 (January 1, 2017): 128–51, https://doi.org/10.1016/j.arcontrol.2017.03.005.tse

[15] Chris M.J. Tampere and L. H. Immers, “An Extended Kalman Filter Application for Traffic State Estimation Using CTM with Implicit Mode Switching and Dynamic Parameters,” in *2007 IEEE Intelligent Transportation Systems Conference*, 2007, 209–16, https://doi.org/10.1109/ITSC.2007.4357755.traffic state

```
      "value": "urn:ngsi-v2:cot-imec-be:device:proximus-100mE39297N31371",
      "metadata": {}
    },
    "binId": {
      "type": "Text",
      "value": "100mE39297N31371",
      "metadata": {}
    },
    "areaCovered": {
      "type": "geo:json",
      "value": {
        "type": "Polygon",
        "coordinates": [
          [
            [4.39415742, 51.20980217],
            [4.39558489, 51.20987106],
            [4.39547565, 51.21076687],
            [4.39404815, 51.21069799],
            [4.39415742, 51.20980217]
          ]
        ]
      },
      "metadata": {}
    },
    "intensity": {
      "type": "Number",
      "value": 53,
      "metadata": {}
    },
    "area_covered": {
      "type": "geo:json",
      "value": {
        "type": "Polygon",
        "coordinates": [
          [
            [4.39415742, 51.20980217],
            [4.39558489, 51.20987106],
            [4.39547565, 51.21076687],
            [4.39404815, 51.21069799],
            [4.39415742, 51.20980217]
          ]
        ]
      },
      "metadata": {}
    },
    "dataSource": {
      "type": "Text",
      "value": "proximus",
      "metadata": {}
    },
    "dateObserved": {
      "type": "DateTime",
      "value": "2020-07-10T14:50:01.158Z",
      "metadata": {}
    },
    "location": {
      "type": "geo:json",
      "value": {
        "type": "Point",
        "coordinates": [4.39481653, 51.21028452]
      },
      "metadata": {}
    },
    "source": {
      "type": "Text",
      "value": "Antwerpen_slimme_zone",
      "metadata": {}
    },
```

```
  "dateObservedFrom": {
    "type": "DateTime",
    "value": "2020-07-10T14:45:01.158Z",
    "metadata": {}
  },
  "dateObservedTo": {
    "type": "DateTime",
    "value": "2020-07-10T14:50:01.158Z",
    "metadata": {}
  }
}
```


Model summary steps

Step 1: Street Cutting

The street cutting is a process where the geolocation metadata of all the data sources is used to divide the given street grid in a finer mesh. The result is a street grid where each segment knows which data points are applicable on it.

Step 2: Transformation of data source types

Data sources of cumulative data types are transformed to the same, directly comparable format. We aim to use data standards such as NGSI as much as possible.

Step 3: Fusion and flow calculation

The data sources are fed as constraints to the cityflows model. The model takes all metadata as constraints into account and calculates a density of people in the street segments. The steps are linked between timesteps.

Requirement parameters are the total estimated number of road users, the number of users on a road segment in the previous time interval, the number of people entering an intersection and leaving it towards another street segment (i.e., continuity constraint). The idea is that for each time interval the total number of people should be distributed over the street segment grid taking into account these constraints.

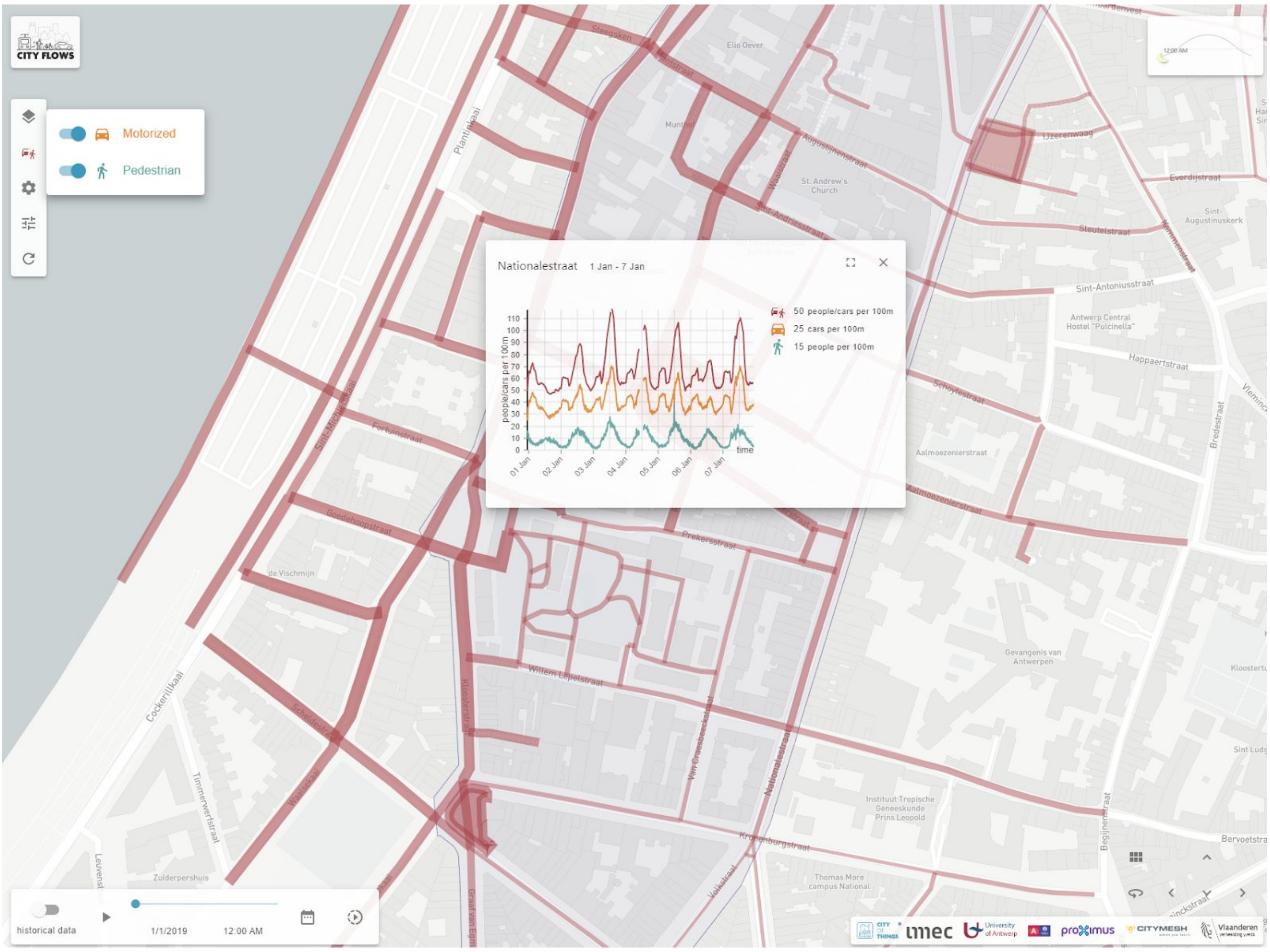
In order to deal with data insufficiency, the gaps are filled in as follows:

- the model takes statistics (for example global split) into account by enforcing that on a global setting the modal split must be 70/30 for example. If local or even dynamic modal split data is available, this can be used.
- in case there is few data, the main model distributes the data from the available data point evenly along the streets, possibly taking into account features such as street with, road width, etc.
- if there is more data, this serves to apply local corrections to the distribution proposed by the model in step 2
- If there is no data source for a certain point, either data can be estimated from in- / outflows of other cells, or no calculation can be done whatsoever.

Outputs

The output of the algorithm consists of a data set giving people densities, for every street segment of the given grid (provided that there is data covering it), split up between the different modalities available from the data sources, see figure 22.

Figure 22: A sample output visualisation



3.1.3.2 Scenarios and Use in DUET

Understanding traffic models and determine optimal sensor placement

The Cityflows model allows to

- compare with averages/ predictions from the traffic assignment models
- understand where additional sensors may be necessary
- reveal that at some locations additional sensors are pointless as they are already sufficiently described by near-by sensors and the constraints they impose.

Adaptive Logistics & Retail

By combining Cityflows output data with other available data sources such as OpenStreetMaps and weather API, we aim to investigate the correlations between flows in the city and the commercial areas in the cities.

Other use cases are still under investigation and will be reported once they have been refined. A preliminary list of ideas and approaches can be found here.

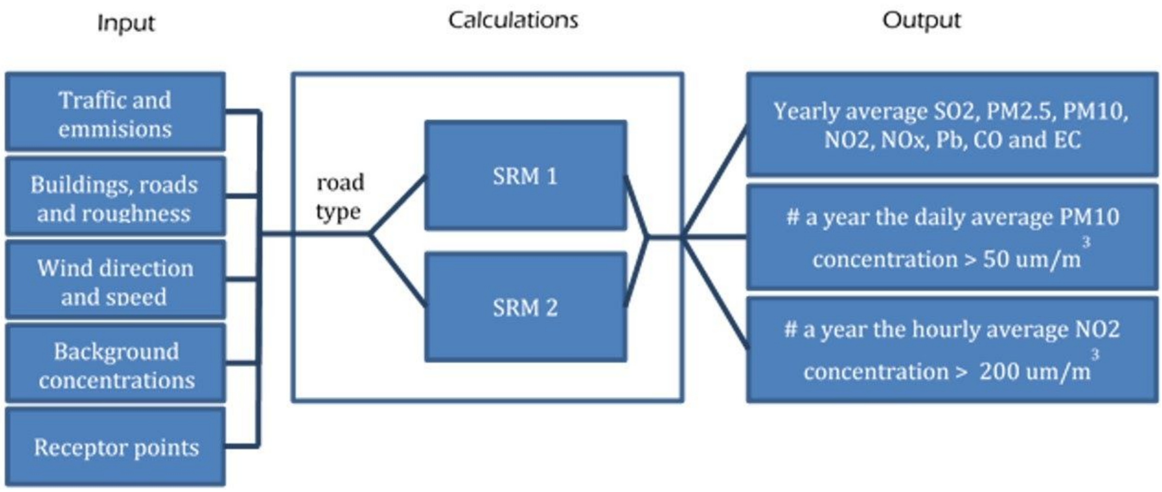
3.2. Air quality emission model

3.2.1. TNO model

3.2.1.1 Introduction

The TNO Urban Strategy model calculates the NO_2 and PM_{10} concentration emitted by traffic. It uses the Dutch SRM1 and SRM2 (Standaard RekenMethode / standardized calculation methods) as described in the RBL 2007 (Dutch regulation for air quality). More information can be found on this website (in Dutch). In this chapter the features and operation, calculation methods, required input and the output of the model is described. A summary of this is shown in the figure below.

Figure 23: Overview of the input calculations and output of the Urban Strategy air module



3.2.1.2 Features and operation

Air quality in cities is substantially determined by emissions of air pollutants by road traffic vehicles. Three components that are usually quantified in order to indicate local air quality are particulate matter (PM_{10}), ultrafine particulate matter ($PM_{2.5}$) and nitrogen dioxide. For PM_{10}, $PM_{2.5}$, as well as NO_2, EU limit values are set for yearly average concentrations, as well as for 8 hour and 24 hour averages respectively.

Dispersion is described using three different sources:

- Background levels are derived from the GCN (Grootschalige Concentratie Nederland): background concentrations, meteo data and terrain roughness.
- Local contribution by traffic using either:
  - SRM2 Gaussian plume model to calculate dispersion of emission from motorways and major roads, or:
  - SRM1 CAR model to calculate concentrations in streets (with buildings along the road)

The contribution calculated by each model is summed for each receptor point and converted to maps for the presentation.

The Urban Strategy air module uses the same standard approach (described above) as other air modules from the Netherlands and that of other countries in the EU, that are specifically used for policy making. It is suited to give an estimation of street values for simple standard geometries. In situations where complex geometries are involved that influence the air flow substantially, this model is not suited.

SRM-1 and SRM-2 models use the same vehicle categories (light, medium and heavy traffic), but SRM-1 has more speed and congestion classes[16]. These are used to select the appropriate emission factor that is input for the dispersion model. Dispersion is calculated across a longer range in SRM-2 than in the SRM1. Also local meteorological profiles are used in SRM-2 and not in SRM-1[17].

The two models rely on solid data on current and future emission estimations of the fleet (how much substance is emitted by a specific type of vehicle), as well as good estimates of background concentration. Emissions and background concentrations change over time. Because of type approval regulations and the introduction of new technologies, like PM filters and automatic driving, emission values are becoming smaller. Because of this, and similar actions on other emitters, background values are likely to decrease as well. Therefore, emission and background scenarios are implemented. On the basis of the year in which developments are planned, appropriate emission factors and background values are taken from the database. These emission and background values are updated annually. For the Netherlands, TNO is the standard for producing the emission factors and the RIVM the background information (GCN). The emission factors are based on emission measurements of the current fleet and expected future fleet compositions. Background information is a combination of model and measurement. For more information on the emission factors see this website and for more information on the background data GCN see this website.

The SRM-1 and SRM-2 are regulated calculation methods, that means that for policy making these models have to be used in the Netherlands and no changes can be made to the calculation methods. TNO is however developing a data driven air module based on SRM-1 and SRM-2, that uses the measurement to auto calibrate the model for the current situation.

3.2.1.3 Calculation method

SRM-1 is intended for calculating concentrations of air pollutants near traffic roads in urban areas, also indicated as "city roads". Characteristic of these roads is that buildings are located in the immediate vicinity, within a few tens of meters of the road. Air vortices around these buildings influence the air flow in the streets and thus the height of the concentrations of air pollution. This is different to highways and other rural roads where the air pollution emitted by traffic does not "get stuck" between existing obstacles, but is directly carried away by the wind. Standard calculation method 2 (SRM-2) applies to this type of road[18].

SRM-1 calculation[19]

When applying this method, the situation considered meets the following conditions:

1. the road is in an urban environment

[16] Smit, R., Mieghem, Hensema, Rabé, Eijk, “VERSIT+ Emissiefactoren voor Standaardrekenmethode 1 (CAR II)”, TNO rapport MON-RPT-033-DTS-2007-00709, 2007, [in Dutch];

[17] Smit, R., Mieghem, Hensema, “Algemene PM10, NOx en NO2 Emissiefactoren voor Nederlandse Snelwegen”, TNO-rapport 06.OR.PT.029.1/RS, 2006 [in Dutch];

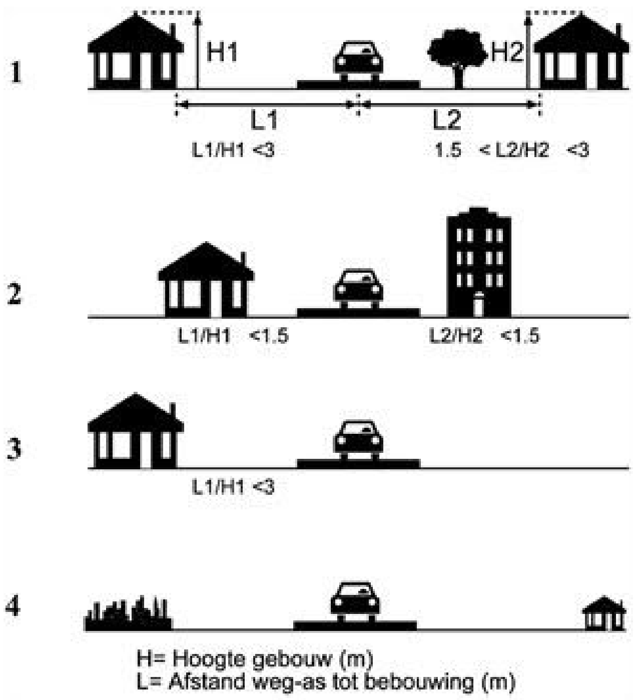
[18] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014

[19] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014

2. the maximum calculation distance is the distance to the buildings, with a maximum of 30 or 60 meters from the road axis, depending on the street type;
3. there is little difference in height between the road and the surroundings;
4. there are no shielding structures along the road.

SRM-1 distinguishes four categories of buildings within 60 meters of the road, see figure 24 (Different road types used for SRM-1). A category with buildings on both sides of the road and less or more connected facades. The same situation, but with relatively high facades, also known as "street canyon". A category with buildings on one side of the road, also with a less continuous wall. In the latter category, the existing buildings are spread around the area, for example a road with semi-detached houses or detached houses. In addition, SRM-1 offers possibilities to take into account any existing trees, a wide central reservation or an exit from a tunnel tube. SRM-1 is not suitable for complex situations with multiple road sections, such as a traffic roundabout[20] .

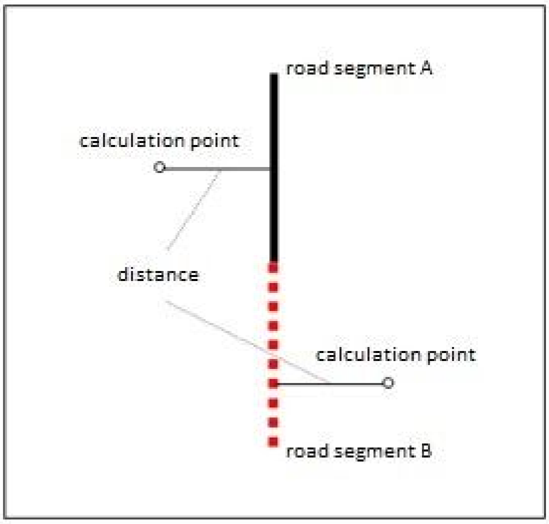
Figure 24: Different road types used for SRM1



The road type is automatically determined based on the geometry LOD 1.1 (outline and the height) of the buildings in the surroundings. In the SRM-1 calculation the concentration is calculated at receptors up to 30 to 60 meters perpendicular to the road and with intervals of 10 meters parallel to the road (see figure 25) the concentration can be compared with limit values from 10 meters onwards, unless the distance between the roads and buildings is shorter. In that case the concentration can be compared with limit values on the facade of the building.

Figure 25: Concentration calculated perpendicular from road[21]

[20] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014
[21] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014



The calculation model makes it possible to perform calculations of:

1. the annual average concentrations of sulphur dioxide, nitrogen dioxide, nitrogen oxides, particulate matter ($PM_{2.5}$ and PM_{10}), lead, carbon monoxide, carbon black and benzene;
2. the number of times per year that the twenty-four-hour average concentration of particulate matter (PM_{10}) exceeds the limit value of 50 µg / m3;
3. the number of times per year that the twenty-four-hour average sulphur dioxide concentration exceeds the limit value of 125 µg / m3;
4. the number of times per year that the hourly average nitrogen dioxide concentration exceeds the limit value of 200 µg / m3.

For an elaborate explanation of the formulas used in these four calculations for SRM-1 see[22].

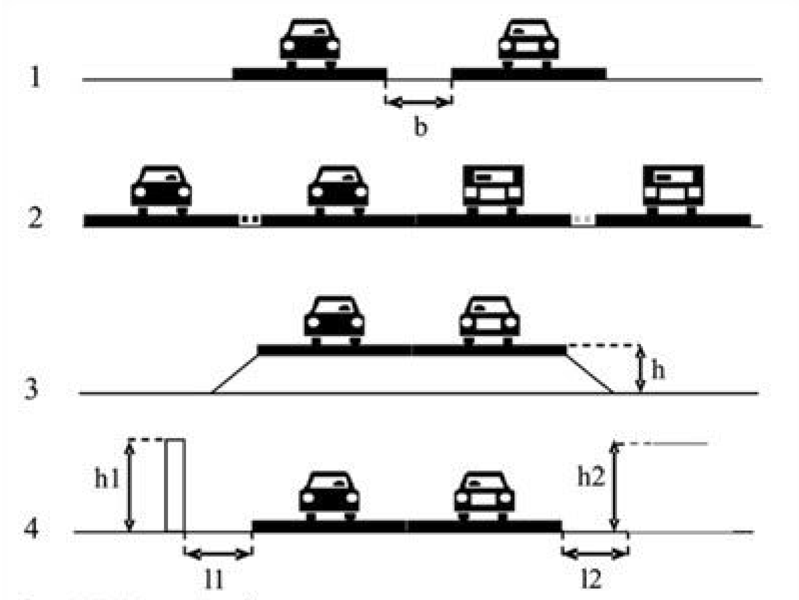
SRM-2 calculation[23]

SRM-2 is a Gaussian plume model and is intended for calculating concentrations of pollutants in the open air in the vicinity of mainly highways and rural roads. When applying this method, the situation considered meets the following conditions, see figure 26 for a reference on the conditions:

1. the presence and width (b) of a central reservation;
2. the configuration of the carriageways. The following configurations are possible:
    a. one direction of travel, consisting of one or more lanes;
    b. two driving directions, consisting of one or more runways;
3. the height (h) of the road in relation to ground level;
4. the presence of screens or ramparts, the location (one-sided / two-sided), the height (h1 or h2), and the distance (l1 or l2) to the road edge, where h has a minimum value of 1 meter and a maximum value 6 meters, and for l a maximum value of 50 meters;
5. the presence of a tunnel, whereby there are no openings in the top or sides of the tunnel.

[22] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014
[23] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 2 (SRM-2) voor luchtkwaliteitsberekeningen, 2014

Figure 26: the different road variants



Calculation method 2 makes it possible to perform calculations of:

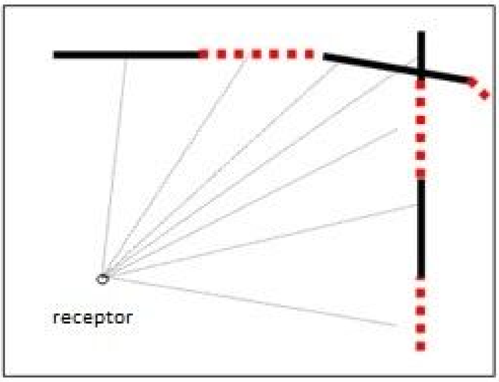
1. the annual average concentrations for sulfur dioxide, nitrogen dioxide, nitrogen oxide, particulate matter ($PM_{2.5}$ and PM_{10}), lead and carbon monoxide;
2. the number of times per year that the twenty-four-hour average concentration of PM_{10} exceeds the limit value of 50 µg / m3;
3. the number of times the twenty-four-hour average sulfur dioxide concentration exceeds the limit value of 125 µg / m3;
4. the number of times the hourly average nitrogen dioxide concentration exceeds the limit value of 200 µg / m3

For an elaborate explanation of the formulas used in these five calculations for SRM-2 see [24] .

The SRM-2 calculations add the contributions of all road sections in a radius of 5km around the receptor, see figure 27. The receptors are placed in a global 10x10 meter grid. The concentrations can be compared with limit values from 10 meters perpendicular to the road.

[24] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 2 (SRM-2) voor luchtkwaliteitsberekeningen, 2014

Figure 27: roads and receptor combination in a 5km radius



Input data

The table below describes the input parameters the SRM-1 and SRM-2 calculations require to perform the calculations.

Table 3: Required input information for the Urban Strategy air module

	SRM-1	SRM-2	Impact
Intensity per type of vehicle Light, medium, heavy	The intensity of light vehicles (passenger cars), medium-heavy vehicles (light trucks), heavy vehicles (heavy trucks), and busses per 24 hours.		Intensity for different classes of vehicles used in combination with the emission factors for the specific class
Road type	May be entered 1, 2, 3, 4. See figure 2 for more details If not entered this is defined automatically, a building shapefile with the attribute height is then needed	Indicate whether the road section is type: ● 3 is a highway. ● 4 is a highway with limited speed by trajectory control.	Has an effect on the air flow and thus dispersion of the substance

Speed (categories)	The speeds has a value **b** to **e** ● **b**, non-urban road, corresponds to roads with average speeds of 60 km/h and with on average 0,2 stops per kilometre. ● **c**, normal city traffic, corresponds to traffic in the city with some congestion. The average speed is between 15 and 30 km/h and with on average 2 stops per kilometre. ● **d**, stagnant city traffic, corresponds to roads at which traffic has an average speed smaller than 15 km/h and with on average 10 stops per kilometre. ● **e**, fast city traffic, corresponds to city traffic with free flow behaviour. The average speed is between 30 and 45 km/h and on average there are 1,5 stops per kilometre. Other speed will fit to the nearest above.	Obligated to fill in: ● Light vehicle 80 km/h, (model knows heavy truck 80 km/h) ● Light vehicle 120 km/h, (model knows heavy truck 90 km/h) ● Light vehicle 130 km/h, (model knows heavy truck 90 km/h) ● Other speed will fit to the nearest above	Speed categories are used in combination with the emissions, In SRM-1 also the city congestion is taken into account. The value will be used to select the corresponding emission factor.
Congestion	Factor that corrects for part of the day with congestion, it is expressed in a percentage, so in case of 7% stagnation, the value is 0.07. For traffic with type **d**, stagnant city traffic, it is already included, the value should be 0.		Used to correct the standard emission to take into account a certain amount of congestion.
Tunnel factor	For each road section needs to be indicated, if it is connected or in a tunnel: ● **0** for tunnel sections ● **6** for sections which tunnel ends		Used to simulate the tunnel end emission. At the tunnel mouth the emission is much larger than on standard road.
Road height	Average height of the road in comparison to the surroundings; a value between -30 and 30 meter.		Has an effect on the range and thus the dispersion of the substance

Screen distance	Distance (in meter) between the middle of the road and the (noise)screen. Up to 50m it affects the calculations.	Forms a barrier for the dispersion of the substance. Effect is dependent on the distance from the road
Screen height	Height (in meter) of the (noise) screen at the left side of the road. Maximum is 6 meter.	Forms a barrier for the dispersion of the substance. Effect is dependent on the height of the screen.

Fixed data

Some of the required data for the calculations are fixed:

- Emission data of the road traffic. This data is necessary to determine the emission caused by different vehicle types
- Background/Double Count concentrations. This data contains all the emission caused by other sources. The amount of emission caused by traffic is labelled in the Double Count data.
- Wind compass rose
- Wind speed (1km x 1km grid)

Emission of road traffic

The emission factors used depend on the vehicle types and speed types and amount of congestion. The emission factor is the amount of emission per vehicle. See table 3 for more information about this data. The emission factors are updated on a yearly basis.

Background / Double-count concentrations

In the Urban Strategy Air Module concentrations near roads are a combination of the emissions of traffic and background concentration. This background concentration (based on Dutch Grootschalige Concentraties Nederland (GCN)) includes the emission of sources like industry, farming, households and traffic. Since traffic emissions already contribute to the background concentration, it is corrected by removing the double counted contributions.

Wind direction effects: (SRM-2 calculations)

In addition to the direct NO_2 emissions, there is also a part that is formed by the reaction of NO with Ozone. The Ozone concentrations depend on the wind direction. For the correction of the amount of NO_2 as a result of the conversion by O_3, there are still the so-called Ozone wind roses.

The Ozone wind rose consists of the map of the Netherlands divided into sections of 1 km X 1 km. For each square, for 12 wind directions (345 - 15, 15 - 45, 45 - 65, etc.) is the frequency of occurrence of this wind direction, the average wind speed and the average O_3 concentration per direction.

Wind speed effects (SRM-1 calculations)

The annual average concentration contribution at a certain location next to a road depends on the average wind speed at this location. For this a wind speed map is used. This wind speed map is derived from actual measurement stations. For the west part of The Netherlands two measurement stations are used to create a wind map. This is done by generating a 1km X 1km map by interpolating the wind speeds of the two measurement stations Schiphol (North-West) and Eindhoven (South-East). This approach can be done with other or more measurement stations.

Fixed data source

For The Netherlands, most of the data is published and updated every year. For calculations in other countries, TNO has an offline Air Quality model LOTOS-EUROS, which can be used to generate the background data outside of The Netherlands, this model can deliver a grid 5x5 km of background concentrations. The model can be used in combination with measurements from background stations to make it more accurate.

Output

An overview of the output of the calculation methods SRM-1 and SRM-2 is given in the paragraph calculation method. The following Figures contain several sample visualisations of the output of the Urban Strategy air Module.

Figure 28: NO_2 concentration map

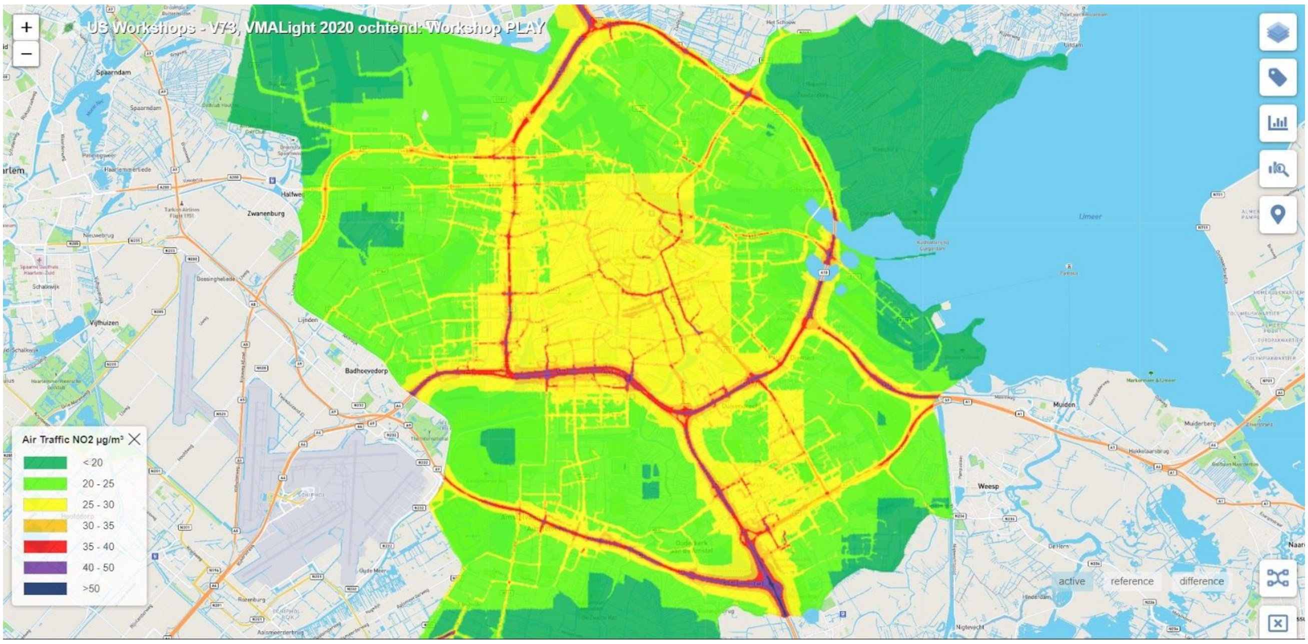


Figure 29: PM_{10} concentration map

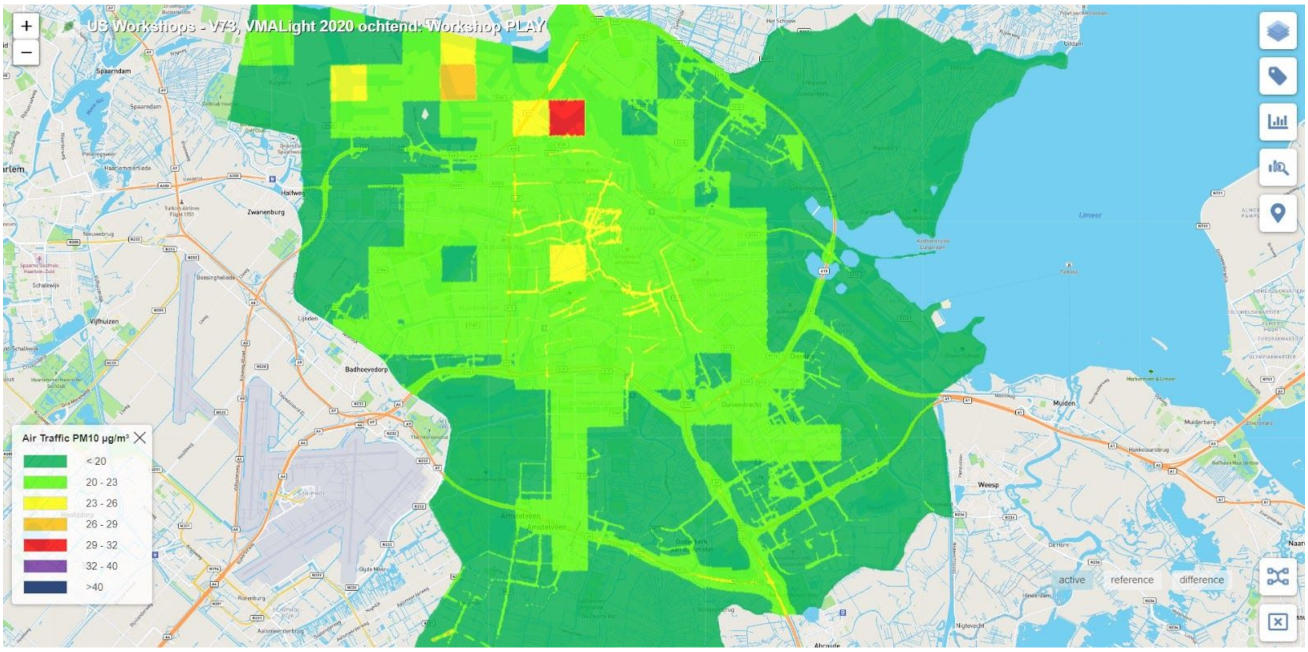


Figure 30: $PM_{2.5}$ concentration map

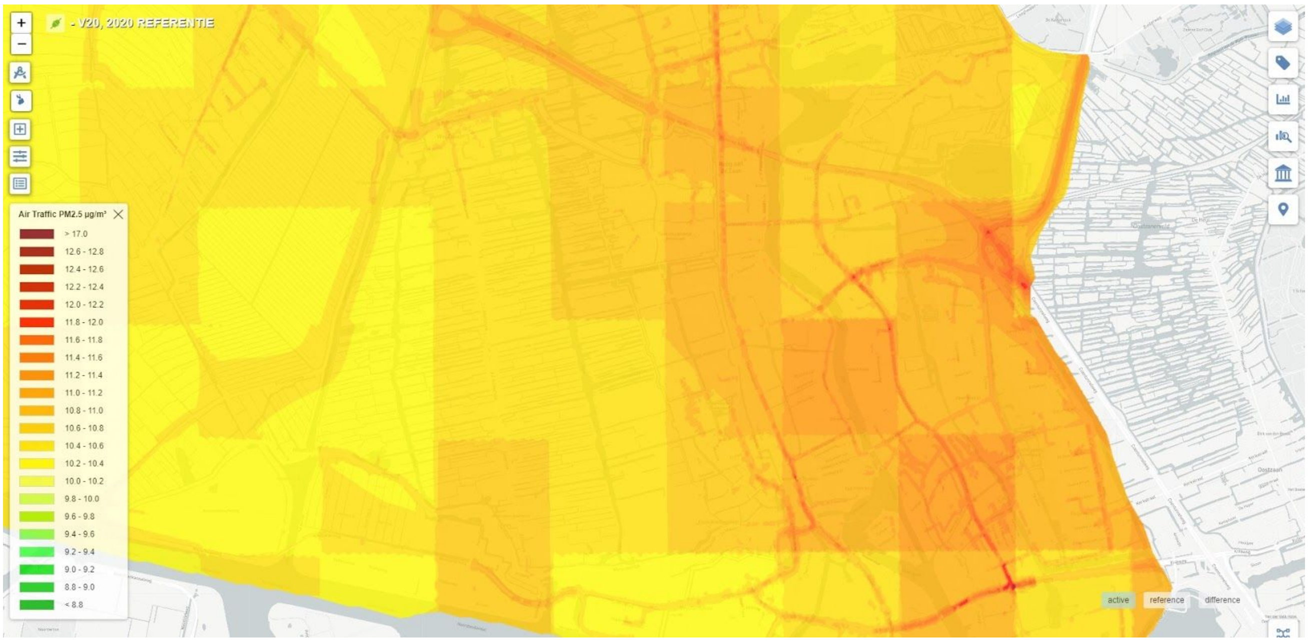


Figure 31: difference plot for NO_2 that shows the effect of a closure of a road in the city centre

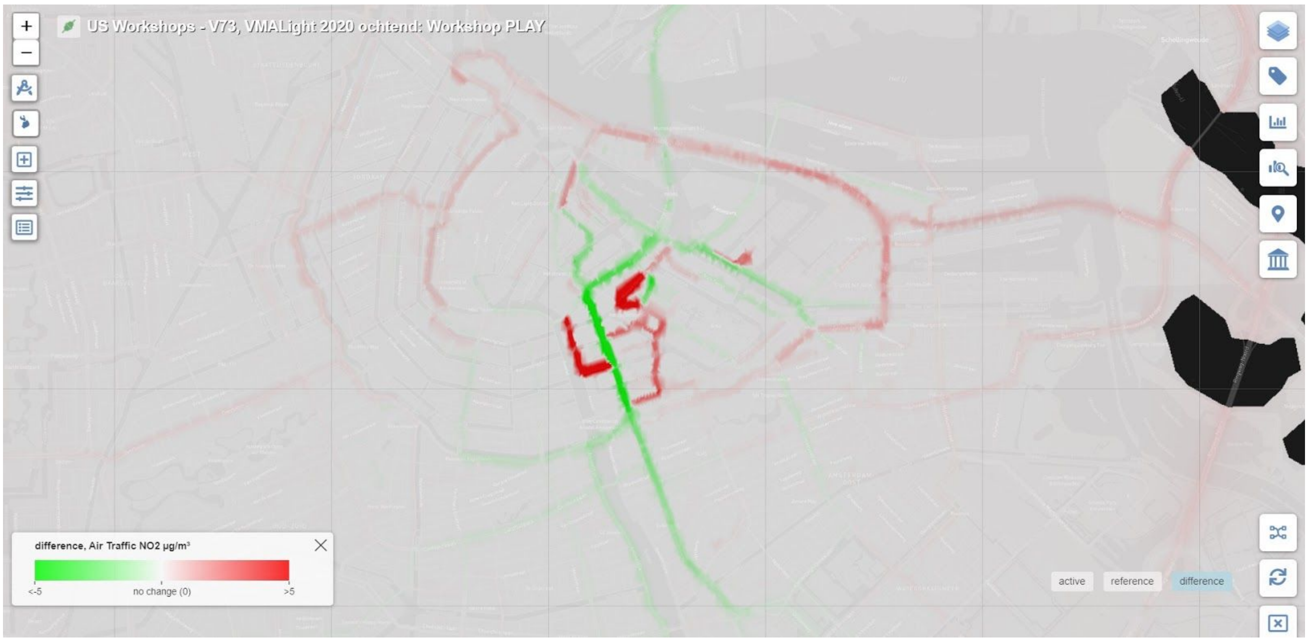
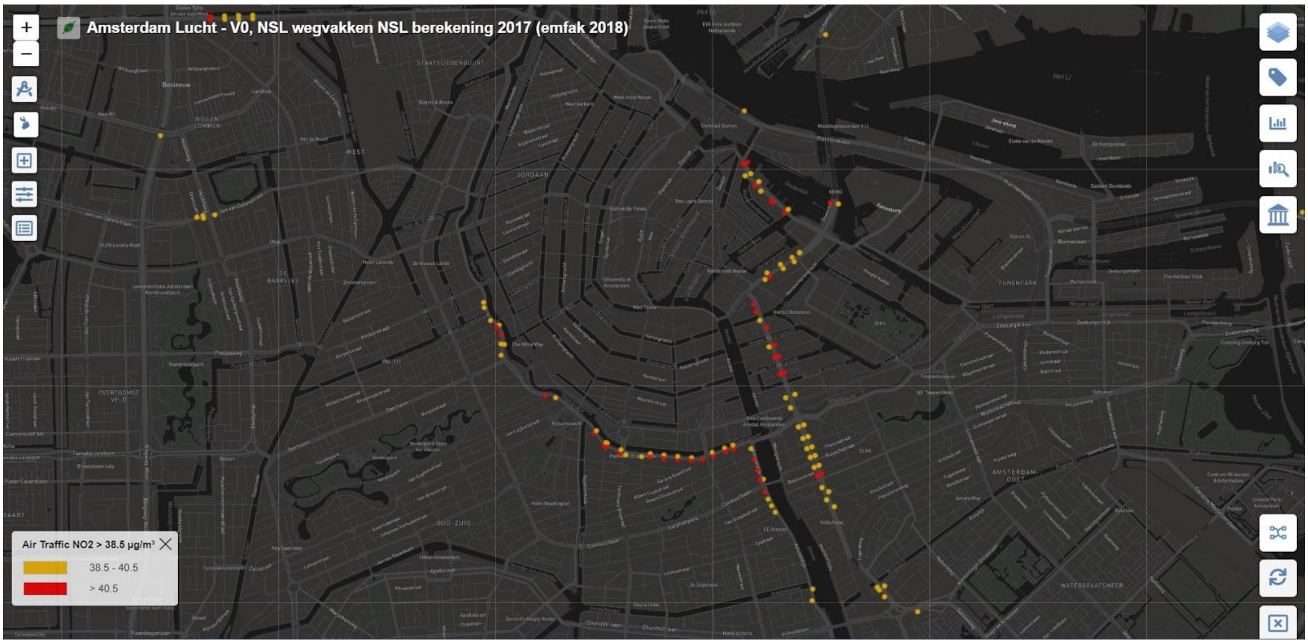


Figure 32: NO_2 exceedances locations



Validity

At the request of the Ministry of Infrastructure and the Environment, the RIVM (National Institute for Public Health and the Environment) compared more than 400 measurements of NO_2 concentrations in 2010 and

2011 with (with SRM-1 and SRM-2) calculated concentrations in 2013. It was concluded that the uncertainty in a single calculation is significant. On average, however, the calculated concentrations are close to the measured values. The similarity is so good that the differences are smaller on average than the uncertainties therein[25] .

The match between the measured and calculated NO_2 concentrations at SRM-2 locations is not bad, certainly not at the higher concentrations. At those higher concentrations, above 40 µg / m3, there is a good match between measured and calculated concentrations. At 40 µg / m3, there is an average underestimation by the calculation method of 0.9 µg / m3. This difference is not significant. However, the number of measurements available at SRM-2 locations in urban areas is limited and the spread and uncertainty are large. These locations have been analysed separately. On average, the results of SRM-2 in these complex situations are satisfactory. At the limit value, the calculations underestimate the measurements in this case by approximately one and a half µg / m3. This difference is not statistically significant[26].

The accuracy of the model is however highly related to input of the model. It requires good validated input data on traffic flows/speeds emissions, background, etc..

3.2.1.4 Scenarios and Use in DUET

The Air Quality model emissions are calculated based on the Traffic Volume on the road network. Other (real time) data such as wind speed & wind direction will also be input for these environmental models.

The Air quality model will calculate the dispersion of air pollution caused by traffic for a grid of geospatial placed calculation points. The results will be converted to map images using interpolation or heatmap technology and placed on top of a map. Calculations are done for several compounds (NOx, PM10, PM2.5, EC, etc.), based on weather information (wind direction, wind speed) and spatial conditions (street canyons, shielding, etc.).

Based on the DUET scenarios, modelling and simulations can provide insight of the impact of policy decisions on Air Quality. Based on certain traffic scenarios Air Quality can be predicted and this will support policymakers and area developers in making complex decisions about urban quality of life, without additional infrastructure changes, or expenditure for cities.

3.2.2. VITO model

The Flemish Institute for Technological Research (VITO) delivers with ATMO-Street and ATMO-Plan two products for simulating air quality concentrations and support for urban air quality plans. Both products aim at the International market and are interesting candidates to be integrated into the DUET Digital Twin as well.

VITO's ATMO-street model calculates air pollutant concentrations across a region, taking into account the regional and urban background but also capturing so-called street canyon effects into one single high spatial

[25] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 1 (SRM-1), 2014,
van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 2 (SRM-2) voor luchtkwaliteitsberekeningen, 2014
[26] van Velze K, Wesseling J. Technische beschrijving van standaardrekenmethode 2 (SRM-2) voor luchtkwaliteitsberekeningen, 2014

resolution air quality map. With this information, we can estimate the health impact of air quality . Using this information, policymakers can take action to reduce concentration levels at pollution hotspots. The data can also be used in applications that can help us, the citizens, for example, to choose the 'healthiest' route to cycle or walk to work.

The current ATMO-Street air quality maps for Flanders, generated by the Flanders Environment agency (VMM) and the Belgian Interregional Environment Agency (IRCEL) are available from the official VMM website. VITO also introduced the ATMO-Street model in China, India and Central and Eastern Europe.

ATMO-Plan delivers a tailored planning module to help the air quality managers, environmental consultants and city authorities in developing medium to long-term air quality strategies. The ATMO-Plan tool estimates the effects on the local air quality of traffic scenario's (both in terms of traffic volumes as well as fleet composition via e.g. LEZ's), urban development plans or industrial mitigation strategies. ATMO-Plan is currently also delivered to the Slovakian and Hungarian governments.

The DUET consortium is currently discussing with VITO and VMM if and how the VITO models can be used and if and how they can play a role as an example of a third party integrated model in the DUET Digital Twin.

3.3. Noise emission models

3.3.1. Noise model

3.3.1.1 Features and operation

The noise models are still under development at P4ALL, preliminary in- and outputs can be found in the Appendix.

Current status: Our plan is to make use of an already existing open source tool called NoiseModelling, developed by Noise-Planet project. The latest big release was described in [27] in the ISPRS International Journal of Geo-Information. We have tested their solution in a desktop environment and now we plan how to best integrate this tool into Traffic Modeller. The goal is to use NoiseModelling tool without the dependencies on any other desktop SW (OrbisGIS, WPS Builder) and also without any other unnecessary tools (H2GIS database and possibly geoserver).

We are in touch with NoiseModelling developers (we had a first conference call in June) and they seem to be interested in our use case. We have detected the first necessary steps as NoiseModelling uses a different database system (H2GIS) to TraMod's database (PostGis). We are currently looking into the scope of necessary changes in the source code - we see the workload estimate as one of P4A goals for the alpha version.

NoiseModelling tool's features include the estimation of traffic noise emission over the transport network and the calculation of sound levels over a receivers grid issued from the propagation from these noise sources to each receiver. (More in ISPRS paper)

3.3.1.2 Scenarios and Use in DUET

[27] Erwan Bocher et al., "NoiseModelling: An Open Source GIS Based Tool to Produce Environmental Noise Maps," *ISPRS International Journal of Geo-Information* 8, no. 3 (March 2019): 130, https://doi.org/10.3390/ijgi8030130.

As described 3.3.1.1 during DUET Plan4All (TraMod team) would like to integrate NoiseModelling tool into TraMod. With TraMod user is already able to explore different traffic scenarios by changing road parameters (free flow speed and capacity). With the addition of near real-time noise calculation users could thus also explore how planned constrictions influence the road surrounding noise level.

3.3.2. Urban Strategy noise model

3.3.2.1 Introduction

Traffic noise in cities affects the lives and health of a large number of people[28,29,30,31] . To regulate and control the effects of urban traffic noise, the European Commission requires major EU cities to produce noise maps and corresponding noise-exposure distributions of their inhabitants[32]. The Urban Strategy Noise Module focuses on the generation of Noise maps and delivers the output on which noise-exposure distributions can be derived. In the Netherlands it is also used as a policy making tool, in order to see the effect of different planning scenarios for infrastructure, buildings and sound barriers.

3.3.2.2 Features and operation

The Urban Strategy Noise module takes into account three different source types:

- Road traffic
- Rail traffic
- Industry

The calculation follows the Dutch statutory methods for noise calculations. The method for road and rail traffic noise is called ''Standaard RekenMethode 2 (SRM2)'' for Noise. A link to SRM2 for Noise can be found here[33]. The method for industry noise is called ''Handleiding-Heten-en-Rekenen-Industrielawaai'' (HMRI). A link to this can be found here[34].

In general the noise module is used to calculate noise levels on receptors. These receptors are placed evenly within the study area, on the facades of buildings and next to roads. From these receptors a continuous noise map can be generated. This can be done for large areas up to 30 km X 30 km. For such an area normally about 3 million receptors would be used, which can give a detailed noise map on street level.

[28] E. Ohrström, L. Barregård, E. Andersson, A. Skånberg, H. Svensson, and P. Angerheim, "Annoyance due to single and combined sound exposure from railway and road traffic", 2007

[29] H. A. Nijland, S. Hartemink, I. van Kamp, and B. van Wee, "The influence of sensitivity for road traffic noise on residential location: Does it trigger a process of spatial selection?," , 2007

[30] Y. de Kluizenaar, R. T. Gansevoort, H. M. E. Miedema, and P. E. de Jong, "Hypertension and road traffic noise exposure,", 2007

[31] D. Ouis, "Annoyance caused by exposure to road traffic noise: An update,", 2002

[32] European Directive on Environmental Noise, 2002/49/EC, URL: http://ec.europa.eu/environment/noise/home.htm , 2019

[33] Reken- en meetvoorschrift geluid 2012 https://wetten.overheid.nl/BWBR0031722/2020-01-01, 2020

[34] handleiding-meten-en-rekenen-industrielawaai https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/brochures/2011/03/22/handleiding-meten-en-rekenen-industrielawaai/handleiding-meten-en-rekenen-inustrielawaai.pdf , 2011

Calculations with the noise module are in general time consuming. To speed up the calculations, Urban Strategy employs parallel computation on the Nvidia GPU cards. With the utilization of GPUs, calculation times of the Urban Strategy noise module are orders of magnitude smaller than calculation times of conventional noise software.

Calculation methods

The calculation methods for traffic noise and industry noise (SRM2, HMRI) are engineering models based on the following basis acoustic relation:

emission level receptor = emission level source + attenuation terms

Where:

- The emission level receptor is the noise level at a receptor;
- The emission level source represents noise emission by the source;
- The attenuation terms represent attenuation of sound waves during propagation from the source to the receptor.

The attenuation mechanisms include geometrical attenuation, air absorption, ground attenuation, noise barrier screening, and reflection attenuation. The noise levels are a function of the octave band frequency, with octave bands 63-8000 Hz for traffic noise

Results

The results of a noise calculation is the noise level at the receptor. In the case of road or rail traffic noise, the noise level is usually expressed as the day-evening-night level (Lden). In the case of industry noise, the noise level is usually expressed as a 24H value, or the so-called 'etmaalwaarde' (Letm).

For an elaborate description of the calculation methods used in the Noise Modules, please see reference[35] for Noise road and rail traffic and see reference[36] for industry noise.

Receptors

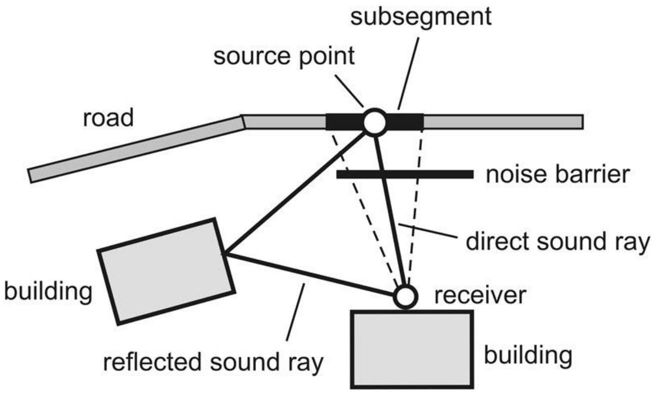
Before a calculation with the noise module can be performed, the set of receptors must be specified. This can be done by uploading a file with receptor locations.

For efficient calculation of continuous noise maps, however, Urban Strategy has a separate module (the receptor module) that produces receptor locations suitable for generating noise maps. The receptors are placed in an efficient way along roads, along facades, and in open areas. The corresponding receptors are called road receptors, facade receptors and open area receptors. This results in a good quality noise map around buildings/screens and in the vicinity of the roads; the receptors are placed on locations where they are needed. Additionally, the facade receptors can be used to create a noise exposure chart; by combining the facade noise levels with the number of inhabitants in a specific building, we can calculate how many inhabitants are affected by a certain level of noise. It is also possible to distribute the inhabitants according to the apartments (including height) and have a facade receptor on each apartment.

[35] Reken- en meetvoorschrift geluid 2012 https://wetten.overheid.nl/BWBR0031722/2020-01-01, 2020

[36] handleiding-meten-en-rekenen-industrielawaai https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/brochures/2011/03/22/handleiding-meten-en-rekenen-industrielawaai/handleiding-meten-en-rekenen-inustrielawaai.pdf , 2011

Figure 33: Noise propagation from source to receptor



Input data

In this section, the input data for a calculation with the noise module is specified. The input data is usually provided as Esri shapefiles. For each new calculation, the input files are first uploaded to the Urban Strategy database (Oracle database). The database is the central point from where the model reads the data for a calculation. The remainder of this section specifies the input data; first the common input is specified, next the input for road, rail, and lastly industry noise is specified.

Common data input

Different shape data about buildings, sound screens, surface coverage and receptors are needed to populate the model:

Type	Shape data
Buildings	PolygonShape, Height, Reflection factor
Sound Screens	PolygonShape, Height, Reflection factor, Screen profile
Surface/Coverage	PolygonShape, TOP10NL
Receptors	PointShape, Height. These are the calculation points

Road data input

Polyline Shape:

The table below specifies the input data (shape attributes) for a calculation of road noise. Three types of road vehicles are distinguished (light, medium-heavy, heavy) and three time periods are distinguished (day, evening, night).

Data	Attribute
Road data	• Shape • Volumn24hours • Percent per hour at Day/Evening/Night • Percent personal auto(PA) per hour at Day/Evening/Night • Percent motors(Mo) per hour at Day/Evening/Night • Percent middle heavy cars(MZ) per hour at Day/Evening/Night • Percent heavy cars(ZW) per hour at Day/Evening/Night • Speed of PA, MZ, ZW, Mo at Day/Evening/Night • Height, RoadType, width

Rail data input

Polyline Shape:

The table below describes the input data (shape attributes) for a calculation of rail noise.

Data	Attribute
Rail data	• Shape • Height • bb_type • IntensityD1 ··· IntensityD11 • IntensityE1 ··· IntensityE11 • IntensityN1 ··· IntensityN11 • StopfractionD1···stopfractionD11 • StopfractionE1···stopfractionE11 • StopfractionN1···stopfractionN11 • V_door1···v_door11 • V_stop1···v_stop1

Industry data input

Point Shape:

The table below describes the input data (shape attributes) for a calculation of industry noise.

Data	Attribute
Industry Noise Sources	● Shape ● Ground_level ● Height ● Sources_type ● Angle ● Direction ● LW31···LW8000 ● RED31 ··· RED8000 ● Groupreduction_day/evening/night ● Cwork_day/evening/night ● NRPNTSRC ● MAXDIST ● LENGTH ● Area ● Intensity_D/E/N, ● Velocity ● Delta_x, Delta_y

Output

The output of a noise calculation consists of the calculated noise levels at the receptors. The results are stored in the database, together with metadata for further processing and visualization.

- Object_id = Unique number linkable to receptor type
- I_D = daytime noise level in dB
- I_E = evening noise level in dB
- I_N = night noise level in dB
- I_LDEN = daily noise level in dB
- I_LAEQ = daily noise level in dB, no correction for I_E and I_N

Figure 34 : Noise L_{den} map for rail traffic

Figure 35: Noise L_{den} Map for road traffic

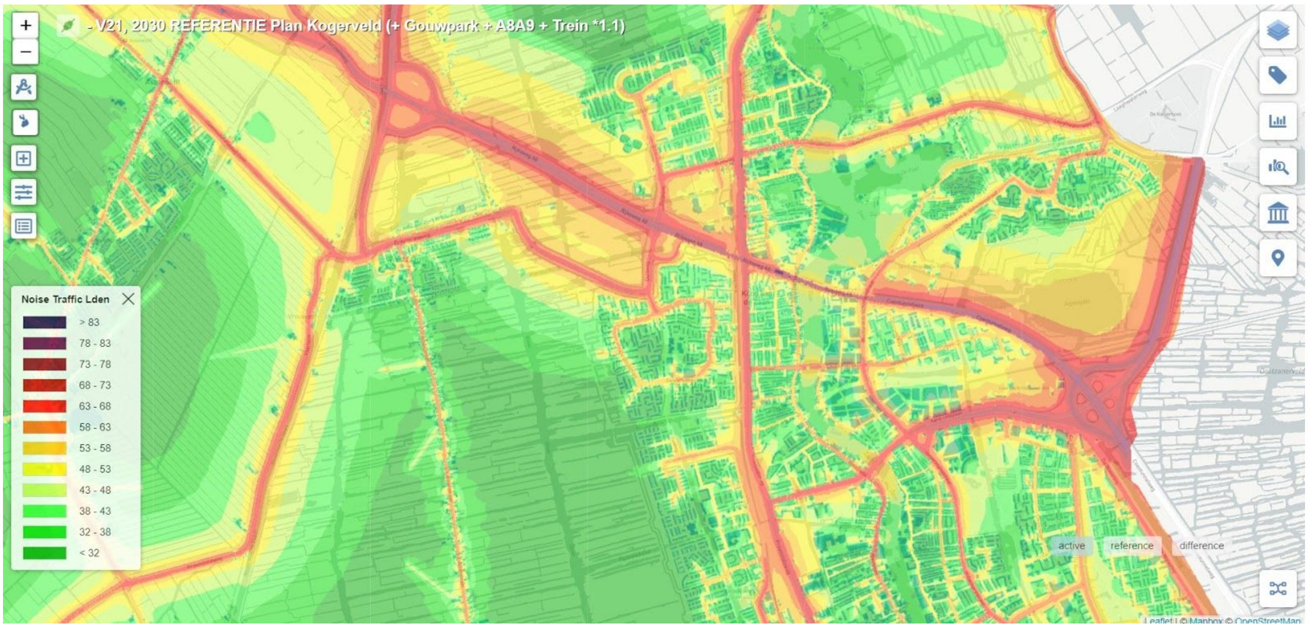


Figure 36: Noise L_den Map for noise Industry

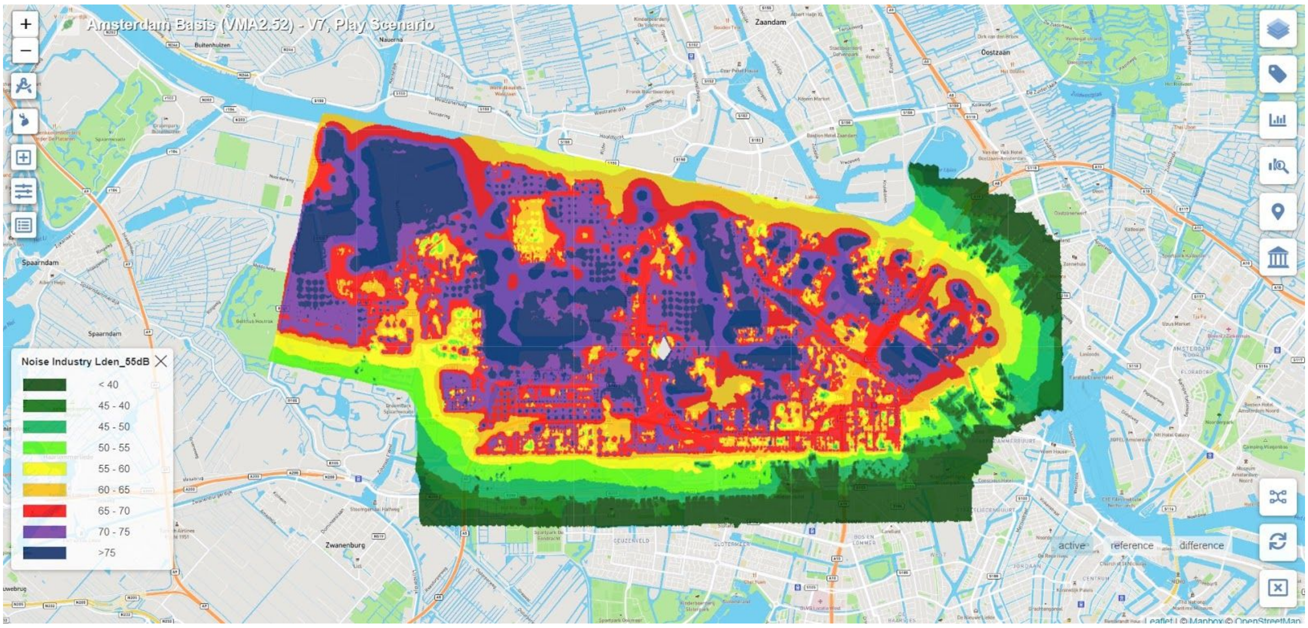


Figure 37: Difference plot for traffic road noise L_{den} that shows the effect of a closure of a road in the city centre

Figure 38: Noise receptor that have levels higher than 48dB(A) noise road L_{den}

3.3.2.3 Scenarios and Use in DUET

The Noise Model calculations are based on the Traffic Volume on the road network.

The Noise model will calculate the propagation of sound caused by traffic for a grid of geospatial placed calculation points. Given the distribution of type of vehicles, noise levels per road link are calculated and the results will be converted to map images using interpolation or heatmap technology and placed on top of a map.

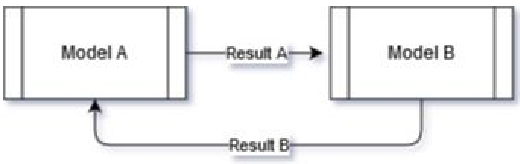
Based on the DUET scenarios, modelling and simulations can provide insight of the impact of policy decisions on Noise Pollution. Based on certain traffic scenarios, noise pollution levels can be predicted and this can support policymakers and area developers in making complex decisions about urban quality of life.

4. Integrating multiple models in a Digital Twin

The way in which models interact with one another depends on how tightly they are intertwined. If you take a look at what is shown in Figure 1 in the Introduction, you see that many of the components are looping back to components that either directly or indirectly were part of their input. Example: The travel demand depends on the generalized cost and the generalized cost depends on the travel demand.

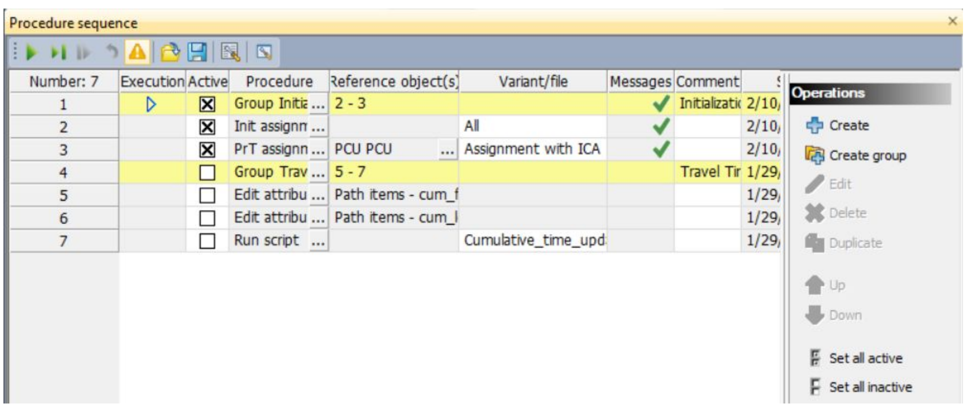
Finding a demand *D* which yields a cost that, if fed back to the demand generation module, gives the same demand *D* is a fixed point problem. You can see this depicted conceptually in figure 39: when results A and B no longer change while running through the loop, we've reached convergence and found a fixed point. In practical applications there will be a check for this condition that has been omitted here.

Figure 39: circular dependencies visualized



Above we call demand generation and cost calculations 'components' or 'modules' because they are the smaller parts that make up what we call a transportation model but they are of course models in their own right.
One of the challenges for DUET will be to find a compromise between decomposition of models into their component models (which allows for more intricate interaction schemes and increases developers flexibility) and the effort involved in providing APIs. A decision needs to be made on what components of larger models should be exposed to the developers using the platform.
It may be necessary to provide the expert user with an opportunity to design their own model-use 'cookbooks' where they indicate calculation sequences of different models for some analysis. This is similar to the procedure sequence PTV uses in their macroscopic modelling Software Visum, see figure 40. Alternatively, we may provide external access to the different modelling tools through an API.

Figure 40: Procedure Sequence GUI



However, the use of these cookbooks will mean that we cannot control what users are requesting from the system, they may define interactions that do not converge at all - that is, to extend the example from above: If Model B depends strongly on the results of A (nonlinearly) and likewise Model A depends strongly on the results of B the loop shown above may never reach a fixed point.
This all sounds very abstract; yet it's nothing more than a modelling equivalent of the behavioural interactions within a city. Entities (companies, individual stakeholders, the general population) adapt their behavior based on policies imposed by the city administration, likewise the administration responds to the

behavior of the actors in the public space. This, of course, is exactly like the kind of circular dependency between A and B that we've introduced above! Therefore, many models that describe interactions between policy makers and stakeholders mathematically require the kind of iterative model computations that are sketched here and in chapter 3 and can be framed within the broader field of game theory. A notable example from the field of transportation that falls into this category are network design problems, they try to answer questions such as:

- Where should we increase or decrease capacities and/ or tolls to improve network conditions and
- what are the system wide optimal traffic signal settings to minimize delays,

a framework addressing any of these questions will utilize a traffic assignment model as a component.
Unfortunately, there exists no generic solution approach that can deal with these Fixed Point Problems efficiently. Obtaining a solution within a reasonable amount of time requires developing approaches tailored to the interacting models, which, at least in the domain of traffic, are often mathematically ill behaved, see[37]. This means that circular dependencies between models within a workflow need to be carefully considered and will usually require the advice of an expert to resolve efficiently.

In the following, we pose some initial ideas for model interactions that may or may not find their way into the final version of DUET, but in any case would serve as useful extensions to the platform. This section is to be elaborated towards v2 and not only show some ideas but present some of the interaction schemes that DUET supports and are integrated into the plattform.

Semi - Automatic Calibration

From a user perspective it's reasonable to expect that the traffic flows predicted by the assignment models somewhat mirror the different observations made by the sensors that are visualized within DUET. An interesting extension to DUET would be a (semi-) automatic calibration module that stores and monitors sensor data and compares this with the flows generated by a calibrated traffic assignment module or/and uses statistical methods to identify sensors that are behaving unexpectedly[38].
If the differences across the network become too large a rerun of the model calibration module may be in order to update the Origin Destination tables. This should be a carefully considered step though, as calibration has a substantial computational cost. These update steps may be triggered automatically or a warning may be issued to the user with a request to recalibrate.

4.1 Low Emission Zone modelling - demonstrating some challenges of integration

The research on integrating models with one another in the smart city modelling domain is sparse and restricted to individual studies rather than a more comprehensive general approach. We want to show here conceptually how traffic assignment models and air quality models can be interfaced with one another to predict the impact of Low Emission Zones. We do this on the one hand to sketch the development and research effort involved in making these model connections possible and valid and on the other hand to demonstrate the effect that erroneous assumptions can have on the outcome of a study.

[37] Srinivas Peeta and Athanasios K. Ziliaskopoulos, "Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future," *Networks and Spatial Economics* 1, no. 3–4 (2001): 233–265.future
[38] This may be due to some external cause such as an accident, social event etc or due to a faulty sensor

Figure 41 was borrowed from [39] and shows the interacting models that were used for an analysis of the impact of a LEZ in the historic city center of Coimbra (Portugal). In order to integrate a similar workflow for a particular city into DUET we would need all the data associated with building a valid traffic and air quality model, as laid out in the previous sections, and information on the fleet composition based on european emission standards. The analysis itself would need to be provided as a separate microservice/model or run as a calculation sequence, see above.

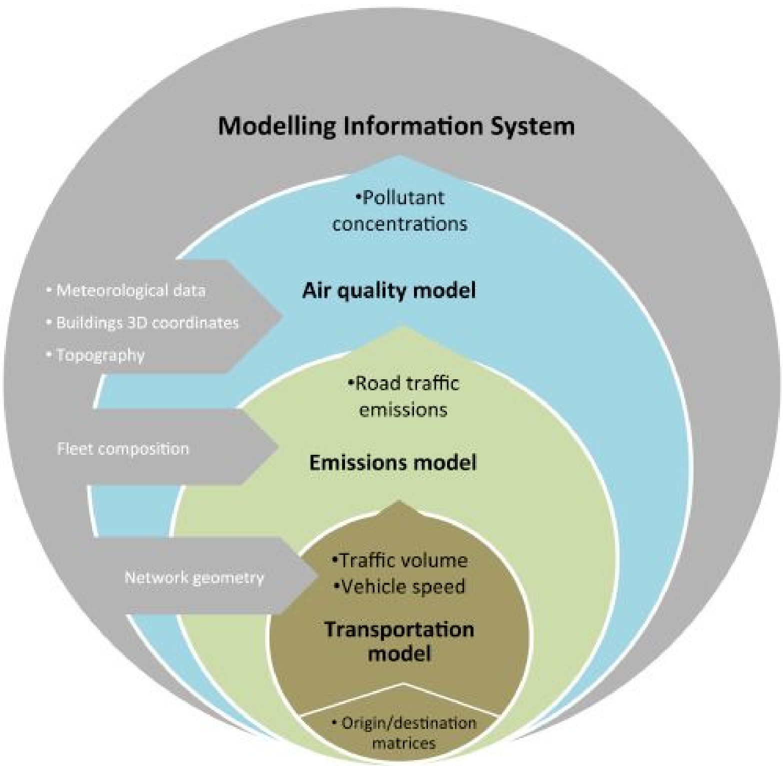
Internally, we now need to keep track of the flow distribution of vehicles among emission classes for each road[40] and make sure that none of the used paths for the high emission vehicles (HEVs) actually cross the emission zone.

An efficient model sequencing to achieve this would be as follows: We first create a copy of the road network without roads that cross the LEZ and calculate a traffic assignment for the HEVs and store the result. We still need to assign the low emission vehicles (LEV) that are allowed to visit the zone. To do so we run an assignment with the LEV trip demand on the whole network taking into account the delays induced by the HEVs. **Note**: The sequencing matters. Had we first assigned the LEVs and then the HEVs we would've not had a converged solution. It would've necessary to recompute the LEVs assignment as the vehicles taking alternative routes circumventing the LEZ may experience a higher cost now and could opt for routes going through the LEZ instead.

Anyhow, the results of the two assignments need to be combined and stored as one assignment result with multi-commodity flow. We can then run the emissions/air quality models and compare the results to some reference scenario.

The emission model itself needs to be able to interpret multi-commodity flow through it's API and assign emissions appropriately.

Figure 41: Modelling chain needed for assessing the impact of LEZ



Interpreting the Results

Back to Coimbra: Dias et al. found that emissions within the zone would decrease drastically with 63% and 52% reductions in PM10 and NO2, respectively.

Outside of the zone however, the model predicted a significant increase in emissions due to the rerouting of vehicles that could no longer travel through the center; as a consequence total emissions in the city increased. This is not unexpected when considering the situation from a traffic viewpoint: by taking away key connections in the city center drivers with HEVs need to make longer detours and ultimately spend more time driving. The size of the effect depends on the network composition in

[39] Daniela Dias, Oxana Tchepel, and António Pais Antunes, "Integrated Modelling Approach for the Evaluation of Low Emission Zones," *Journal of Environmental Management* 177 (July 15, 2016): 253–63, https://doi.org/10.1016/j.jenvman.2016.04.031.

[40] In academic literature this is called multi-commodity flow

the city under consideration. It's reasonable to assume that it will be less severe if there are viable alternative routes circumventing the LEZ with sufficient capacity to not cause new congestion[41]. E.g. if vehicles take detours where they can keep their speed constant without stopping due to congestion or traffic lights their emissions impact may be reduced or at least not increase drastically compared to them travelling through the center.

One of the limitations that were recognized in the study is that possible changes in the composition of the vehicle fleet were not taken into account.

An empirical analysis of german LEZs by Wolff [42] found that they reduced PM10 emissions on average by 9 %. The study considers the changes in fleet composition as one of the major reasons why LEZs do have a positive impact. They observed the most drastic changes in the commercial part of the vehicle fleets.

This information should be taken complementary, on the one hand we learn that traffic and air pollution may deteriorate for a while after installing an LEZ depending on the network composition, on the other hand expected changes in the vehicle fleet need to be assessed either qualitatively or through a vehicle fleet composition model, see [43] for an overview.

A more general conclusion that can be inferred is that use of models in the Digital Twin context should be combined with reading up on studies that assess the impact of the proposed new policy and critically verify whether the used models capture all the effects that empirical analysis has deemed relevant.

[41] Air quality is closely linked to the prevalence of congestion, as can be seen in the model description of the Air Quality models

[42] Hendrik Wolff, "Keep Your Clunker in the Suburb: Low-Emission Zones and Adoption of Green Vehicles," *The Economic Journal* 124, no. 578 (August 1, 2014): F481–512, https://doi.org/10.1111/ecoj.12091.

[43] Jonatan J. Gómez Vilchez and Patrick Jochem, "Simulating Vehicle Fleet Composition: A Review of System Dynamics Models," *Renewable and Sustainable Energy Reviews* 115 (November 1, 2019): 109367, https://doi.org/10.1016/j.rser.2019.109367.

5. Conclusion & Vision towards v2

The document at hand first and foremost gives an overview of the models that the different partners can provide to DUET. The modelling partners (TNO,KUL, P4ALL) have taken different approaches to describing their models, part of v2 will be to make these model specifications more homogeneous in form. Each should contain a conceptual description of the model, more detailed technical requirements/ inputs, some notions on validity and its dependency on data and a link to specific epics.

Meetings leading up to this deliverable have revealed that there is a mismatch between some of the epics that are being pushed and the modelling expertise of the partners. The partners involved with mobility modelling (KUL/P4ALL) use regional traffic models to predict traffic flows between different areas in a given study area, yet many of the epics that are proposed in D2.3 either concern much more local problems such as parking or involve domain models that we do not have readily available in the consortium (Public Transport/ Modal Choice). In the months going forward the modelling partners and pilots need to be involved in a conscious effort to align on feasible epics - both from a modelling - and data perspective.

The introduction of this deliverable addresses these concerns on a more conceptual level and sketches how models operate at different geographical scales. On the one hand, this is closely linked to the effort involved in data acquisition - finer grained models typically need data at a higher spatial resolution - and on the other hand it affects model validity; the functional relationships and parameters that go into a descriptive model are typically developed for a range of geographical scales and may not be valid when 'zooming' in or out too far. For a successful study within the Digital Twin environment the user needs to be aware of such pitfalls and choose the appropriate tools to answer his or her policy inquiries. This deliverable takes a first step in creating a shared understanding of the limitations and data requirements for the use of models in evidence based policy making.

This deliverable also provides a sort of reference for the system architects that need to facilitate the data-flows sketched in the model specifications.

6. Appendix

6.1. Static traffic model (P4ALL)

Data inputs

Full database schema for TraMod input data can be found here.

Traffic generators (mandatory) - demographic data about places that are usually represented as points. These points can be cities, city districts or building blocks – it depends on the granularity of the data and the desired level of detail. These data are used for estimation of traffic flows in the network. Distinguishing between different types of places such as living, industrial, service or shopping place is useful for estimation of traffic flows direction changes in time. For required attributes see database schema.

Road network (mandatory) - well defined and topologically correct road network is the fundamental constraining graph structure, which describes the allowed movements between different places. For required attributes see database schema.

Calibration measurements (possibly optional) - physical measurements of traffic volumes (traffic census) at particular spots of the traffic network are used for calibration of calculated volumes. If there is an already calibrated OD matrix provided there is no need for additional calibration measurements. For required attributes see database schema.

OD matrix (possibly optional) - Origin-destination corresponding with the provided traffic generators. If the attributes provided for traffic generators and calibration measurements are precise, OD matrix can be estimated - however, not in near real time but (usually several days are needed). For required attributes see database schema.

Model outputs

Output from TraMod can be fed to any service using the TraMod API. The API's attributes are described in detail here.

TraMod is a tool for effective traffic management being able to:

- Explore and analyze the past traffic
- Get real-time information about the current traffic
- Be able to model "what-if" analysis of city traffic
    - tactical level ~ planned cultural or sport events
    - strategic level ~ planned roadworks
    -

Limitations - GIGO (garbage in, garbage out), the precision of results strongly depends on input data (road network modelling incl. turns; relevant distribution of Zones, calibration using sensors)

Sources for additional info:
https://trafficmodeller.com/
Source code:

back end: gitlab.com/kolovsky/traffic-modeler
TraMod 2.0 back end: https://gitlab.com/kolovsky/traffic-modeler-rest
API: https://gitlab.com/kolovsky/traffic-modeler-api

P4A - TraMod - Duet - tech survey questions

6.2. Noise model (P4ALL)

Data inputs

Mandatory:

- Roads (source) (line) – Roads layer with convert AADF to annual average day/evening/night flows.
- Buildings (polygon) – Building geometry layer that provide information about *HEIGHT* of buildings.
- Receivers (points) – Regular grid of receivers around buildings with step attribute. Step means step value of the grid in a Cartesian plane in meters.

Optional:

- Ground Type (polygon) – Ground provides geometry about areas like parks, parking places and other public places.

Model outputs

The output from noisemodelling algorithm is receivers layer / column in database, which provide information about noise pollution. This is basically columns with decibel value in the range from 63 to 8000 hertz. That can be visualized in any GIS software or next step is interpolation point layer to GRID file.

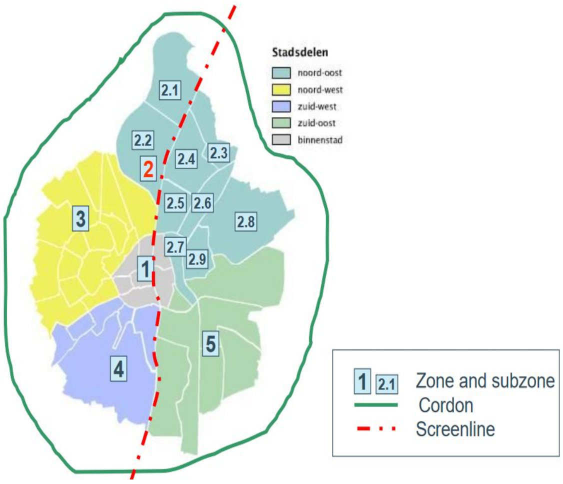
Sources:
PDF documentation (April 21, 2020): NoiseModelling Documentation Release 3.0
about NoiseModelling, see the website
see the documentation
Source code: github.com/Ifsttar/NoiseModelling
See the CNOSSOS-EU traffic model which is implemented in noisemodelling emission model

6.3 Showcasing how a basic Traffic model can be created

In the interest of creating some shared understanding of the amount of effort that is involved in building **valid** traffic models we describe here how a simple unvalidated traffic model can be created from scratch and serve as a starting point for calibration efforts.
In the scope of an educational project at KU Leuven we've started to create a modelling toolkit in python for Static Assignment that will eventually (under development) allow us to create simple and rather crude traffic models for any region.
You can find a demo here , it should make the following text a bit more tangible.
Recall that for a traffic model we need two fundamental inputs, a network graph and travel demand, see I/O schemes in deliverable 3.3 .
The necessary road network graph and travel demand can be generated using OpenStreetMap, a Volunteer based Geographic Information System. One of the packages that we utilize to do so is OSMnx, it does most of the heavy lifting of interacting with the OpenStreetMap API and enables us to extract network graphs in

python's dictionary format, see networkx. For a better understanding of how a network graph is generated based on the raw OSM data we refer to the documentation of OSMnx and the associated publication[44].
We determine capacities of the links based on their road classification and speed as showcased in [45].
We're currently working on incorporating demand generation (what you see in the demo is randomly generated). The idea is to extract Traffic demand based on OpenStreetMap data; the connection here is not entirely obvious. To thoroughly understand it, it's necessary to dive a bit deeper into traffic demand modelling. We sketch this here to convey the complexity of demand modelling and the inherent difficulty of obtaining valid results, which of course motivates the need for calibration.

Traffic demand modelling - a short introduction
In a traffic model the geographical scope is typically divided into the *study area* and the *area of influence*. If you were to make a model for a city, any of the surrounding cities or urban satellites are part of the area of influence. A sizable number of trips may be going into or out of the study area from/to these adjacent settlements.
Figure 42: An example of zoning



The study area and area of influence are delineated into a number of zones, they typically correspond to administrative units or statistical zones to make links with census data easy - but that must not necessarily be so. Depending on the provided socio-economic data and its accuracy smaller zones can be constructed. Figure 42 shows an example of what zoning looks like in practice.

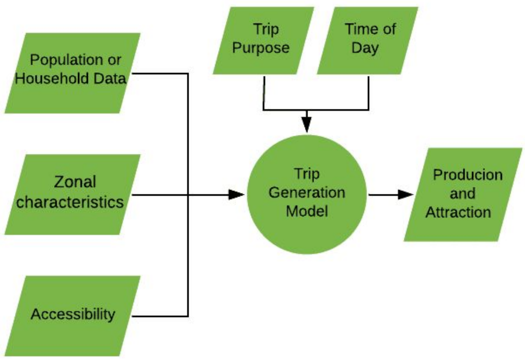
The next step is to determine the production and attraction of all zones, they represent the number of trips originating in- or ending in each zone. The following factors have been taken into account in these models:

- Population Characteristics
  Or/ and Household Characteristics
  - Income
  - Car Ownership
  - Composition (number of people, number of employed people
- Zone characteristics
  - Land use
  - Land price
  - Degree of Urbanization
  - Residential density
- Accessibility
  - Extent of transport options
  - Quality of transport

[44] Boeing, "OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks."
[45] Zilske, Neumann, and Nagel, *OpenStreetMap for Traffic Simulation*.

Regression analysis is often used to calculate the impact that these factors have on production/attraction. Typically production- and attraction models are handled separately and further differentiated by the trip's purpose and time of day, see figure 43. Some of the inputs listed above may be relevant to production models but not attraction models and vice versa or their coefficients in the regression may differ. E.g. It's clear that sqm of industrial area will be correlated with a higher attraction of work related trips in the morning whereas it may have no statistically relevant impact on the number of leisure trips. Likewise work trips may not be as sensitive to the density of the infrastructure of a zone due to their mandatory nature, whereas we expect the impact on leisure trips to be higher.
The distribution of trip purposes varies over the day, most people during the morning peak may be travelling to school or work - whereas during the day the trips that are made are much heterogeneous by purpose.
For coarse grained traffic models zone characteristics and network characteristics as proxies for accessibility may suffice as an input for a rough sketch traffic model. For more refined models with smaller zones household or other more detailed population characteristics are needed.

Figure 43: Trip Generation model I/O scheme



Trip generation models have their own pitfalls and peculiarities, we cannot address them here for the sake of brevity - we refer to [46] for more details and an introduction to category analysis as an alternative to regression analysis.

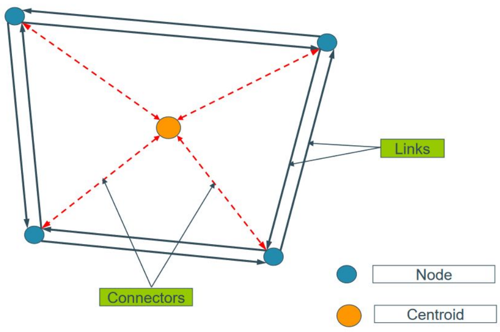
OSM incorporates zonal characteristics and can be used to extract accessibility indicators. Given a calibrated Regression model (which can be found in academic literature) it can be used to get an estimate on production and attraction.
Yet this is just the first layer of the decision pipeline that is used in OD matrix generation. Trip Distribution models match the production and attraction to actual trips and Route Choice models split up OD matrices into different parts per mode. Each of these layers is ideally refined by data and calibrated in some way - although there is not one single right way of doing this. The methodologies that are being used in studies vary on the data that is available and the budget that one is willing to invest into getting it. We're in the process of developing a generic procedure solely based on OpenStreetMap here.
We intend to utilize this approach to generate OD tables for any region based on open data sources.

Shortcomings in zonal modelling of traffic flows

It's noteworthy that traffic inside a zone, that is with its origin and destination as the same zone, is not modelled at all. Furthermore, zone demand is typically represented by artificial nodes called centroids and then connected to the network through artificial links called connectors, see figure 44.
This representation of the beginning and end of the trip is a simplification that has worked well for Static Traffic Assignment but it doesn't allow to make definitive statements about what happens in the origin- or destination zones. This is related to what is sketched in the introduction of D3.3 - models operating and being valid at different geographic scales.

[46] Ortúzar and Willumsen, *Modelling Transport*.

Figure 44: network representations of demand



The more generic message that we're trying to convey here is that models are useful to an extent, they can facilitate analysis and aid decision making - we must keep in mind however that they will always be an imperfect representation of reality which makes enrichment with data and calibration so vital. It's easy to draw the wrong conclusions based on insufficient data or the wrong assumptions if a study is not validated with care.