

Prof.dr.ir. J.W.C. van Lint

Traffic under a Microscope

Inaugural speech

April 2nd, 2014

Faculteit Civiele Techniek
en Geowetenschappen



DiTTLab

Delft integrated
Traffic & Travel
Laboratory

Traffic under a Microscope

Inaugural speech

Short version has been spoken on April 2nd, 2014
at the occasion of his acceptance of the
position of full professor of 'Traffic Simulation & Computing'
at the Faculty of Civiele Techniek en Geowetenschappen
of Delft University of Technology

by

Prof.dr.ir. Hans van Lint

Inaugural speech Prof. Dr. Ir. Hans van Lint
AvL Chair Traffic Simulation & Computing

April 2nd, 2014
18:00 Inaugural speech
18:30 Reception
Aula / Congresstruim
TU Delft

Knowledge discovery in the traffic sciences begins and ends with data, just like in any empirical science.

However, processing data begins and ends with knowledge. Without ideas of what we might find and how we could use it, we're probably not gonna find anything (useful).

I will present my plans for an integrated traffic and travel laboratory that will combine (open) data with (open-source) traffic simulation

Traffic under a Microscope

In this lab we put (big) data in perspective, and ...

DITTLab full integrated traffic & travel laboratory

Past Present Future

Traffic data in infrastructure
Road quality & safety
Traffic data in infrastructure & prediction
Comprehensive knowledge generation
Virtual reality training

Predictions Data Necessities Ideas & Theories

*Mr Rector Magnificus, members of the Executive Board, fellow professors and other members of the university community.
Ladies and Gentlemen,*

When my dean called me in July 2013 to confirm that “our faculty had a new AvL Professor” the news was not unexpected, but I still needed a drink for the reality to sink in. An Antoni van Leeuwenhoek (AvL) chair at TU Delft is a great honour and I am joining an elite group of highly esteemed colleagues. An AvL chair also offers an outstanding opportunity to focus on a subject that does not yet have a dedicated chair, but deserves to have one. In my case, that subject is traffic simulation, including all the data processing and computing needed to make meaningful predictions with these simulation models. To simulate traffic, theories and mathematical models are needed that describe the underlying behaviour of participants in traffic—how they drive, walk or cycle. And that is only half the work—and probably much less than that. These mathematical models then need to be translated in efficient and robust computer code. Traffic simulations also require a large amount of input data and have large numbers of adjustable parameters that determine driving and travel behavior in these simulations. All of these variables and parameters need to be estimated based on data, using mathematical and computational methods. And finally, to view and interpret the results, advanced data processing and visualization tools are needed. Clearly, many disciplines are involved in the development of useful and successful traffic simulation models.

Nonetheless, the starting point for simulating traffic is knowledge about the underlying traffic processes. In this booklet, I will briefly outline what we know about traffic (and how we know it); what we do not yet know; and how I hope to contribute to plugging that gap. But let me start by briefly telling you something about Antoni van Leeuwenhoek himself.

Antoni van Leeuwenhoek (1632-1723) lived much of his life just a stone’s throw away from my own house in the most beautiful city centre in the Netherlands—forgive my bias, but I was born and bred in Delft. Trained as a craftsman in textiles and fabrics, Van Leeuwenhoek was fascinated from an early age by viewing and analysing these materials through magnifying glasses and microscopes. The microscope was probably invented in the Netherlands (that honour goes to Zacharias Jansen in the late sixteenth century) and was made famous by the Englishman Robert Hooke (a contemporary of Van Leeuwenhoek). In cork and other materials, he identified amazingly regular blocks that he described as “cells”. What Hooke saw were not actually cells, but hardened husks around

cells that had long since perished. Van Leeuwenhoek was the first person to actually observe cells, also in living matter.

Another famous son of Delft and a personal friend, Rienier de Graaf, brought Antoni van Leeuwenhoek into contact with the Royal Society, an important European scientific forum that remains so to this day. The result of this introduction was more than 500 publications, in which Leeuwenhoek and his superior microscopes made the microscopic world accessible to science and the wider public for the very first time. Van Leeuwenhoek was the first to describe bacteria and cells of all kinds of plants and animals, including the cells he was able to take from his own body without any permanent damage.

Van Leeuwenhoek is without doubt the Godfather of microbiology, but he also features in the “Top-5 lists of influential scientists” of colleagues from very different disciplines, ranging from cosmology to particle physics and engineering sciences. This is not only because of his many publications, but primarily because of the tools he developed that enabled science to take an enormous leap forward. As a result, every TU Delft professor owes a debt of gratitude to this brilliant local forerunner.

My aim is to follow in Van Leeuwenhoek’s footsteps: develop tools to observe and investigate traffic more effectively. Although it is impossible to achieve even a fraction of the impact that Van Leeuwenhoek had, I have the advantage of being able to build upon the work of, and work in cooperation with experts from many different disciplines, within TU Delft and beyond. A fantastic prospect!

WHAT DO WE KNOW ABOUT ROAD TRAFFIC AND HOW?

The answer to the second question is simple: by looking at data. Figure 1 illustrates the route covered from data to discoveries and ideas, from ideas to theories and via predictions back to data.

- In this case, it starts with video footage taken from a helicopter (Figure 1 upper centre). We translate this video footage into vehicle trajectories, using advanced video-processing algorithms. This kind of trajectory is a line that shows the vehicles’ position over time, and very many trajectories on a motorway segment look like the tangle of spaghetti Figure 1 (upper right).
- The next step is to search these trajectories for recurring patterns. These give us ideas and hypotheses—could it be that...? In this case, one of the hypotheses is that space taken up by a lorry compared to a passenger car

can be expressed as a function of the speed. After all, if the road is not busy and the speed is high, relatively speaking lorries do not take up much more space than passenger cars. However, that changes as the speed reduces. In the extreme case, a traffic jam of 100 lorries is at least four times longer than one with 100 minis Figure 1 (lower right) illustrates this principle. Of course, there are many other differences in driving behaviour and vehicle characteristics between lorries and minis that are of relevance.

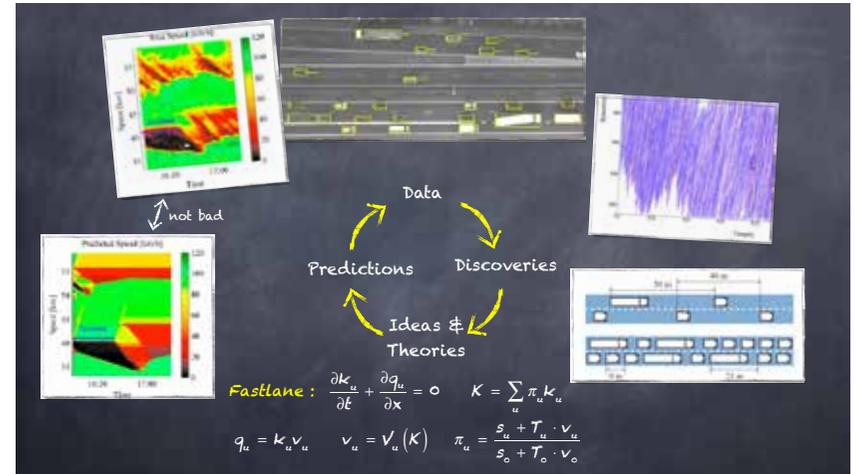


Figure 1

- When enough of these ideas have been developed, they can be structured to form a coherent theory and associated mathematical models. Figure 1 (lower centre) shows the maths behind Fastlane, a new multi-class macroscopic traffic model that can explicitly include different vehicle categories and their different characteristics. I will return to that later.
- Finally, we test (or rule out) whether the model is capable of making predictions in all kinds of relevant conditions. In this case, Fastlane is used to predict traffic dynamics on the A15 motorway in the direction of the port of Rotterdam. It is particularly important to model freight traffic effectively on this motorway. Figure 1 (upper left) shows a so-called heatmap (speeds over space and time) of the real data; the Fastlane approximation is shown on the lower left. In this case, the scenario involves an accident in which one lane had to be closed. The results suggest that this approach is very promising.

So we proceed from data to discoveries in these data; from hypotheses to theories and then via predictions back to data. In a nutshell, that is the scientific method! The circle is continually repeated and each round provides new insights, ideas or theories. The step from a mathematical model to a simulation model that can be used to make predictions is a point I will return to later. But first, a public secret: the circles and arrows diagram in Figure 1 shows a research process that is far too structured. In an witty TED talk, neuroscientist Stuart Firestein describes the research in his lab as follows: "It looks a lot less like the scientific method and a lot more like farting around ... in the dark"ⁱ. In other words, we often mess around with ideas and hunches until we find something that works and is worth pursuing.

However, there is a clear structure in that messing around. Research is unsuccessful much more often than it is successful and the mistakes and dead-ends provide direction and progress. Hypotheses that do not hold up make it clear what we do not yet understand. Research strategies that prove unsuccessful pave the way for research questions that are of relevance. Only then do we identify appropriate experimental methods and the interesting answers that can be used to develop solutions and applications.

This means that research can start anywhere in the circle. Sometimes, the quest starts in the dark, with a fundamental question or a stroke of genius. Think of the famous Higgs boson. At other times, the quest may start with observations and it may last a long time before a theory is identified that explains them. Darwin's inspired idea of evolution through natural selection came about in this way.

A better metaphor for conducting scientific research (than applying the rules of the scientific method in a highly structured way) is playing a game in which neither the result nor the path leading to it are determined in advance. Instead of throwing dice, we use logic and observations to determine the next move. The rules of the game are simple: leave everything open to discussion, test every idea and pursue the empirical evidence, even if it points in exactly the opposite direction than was originally thought. This is why universities need to invest in well-equipped playgrounds in which academics and students can "have fun messing around". The game of research can be played at the highest level in such playgrounds. And if industry and government occasionally join in with the game, it can result in wonderful innovations. I want to build a playground like this for traffic research, because progress in this field is sorely needed. The history of modern traffic research dates back to the 1930s.

ⁱ http://www.ted.com/talks/stuart_firestein_the_pursuit_of_ignorance

Bruce Greenshields (1893, Winfield (Kansas) - 1979, Tyler (Texas)) came up with an ingenious way of testing a slightly older hypothesis back then. The hypothesis is that the speed at which someone drives is proportionate to the distance to the vehicle in front. The shorter the distance gap, the lower the speed. One might also say: the lower the speed, the closer to each other we dare to drive—there's no causality implied here. In order to quantify this statistical relationship, Greenshields used a camera and an ingenious system to take a series of photos of passing cars automatically. With the data he collected, he constructed the first empirical speed-density relationship (known as the fundamental diagram of traffic flow - Figure 2 left).

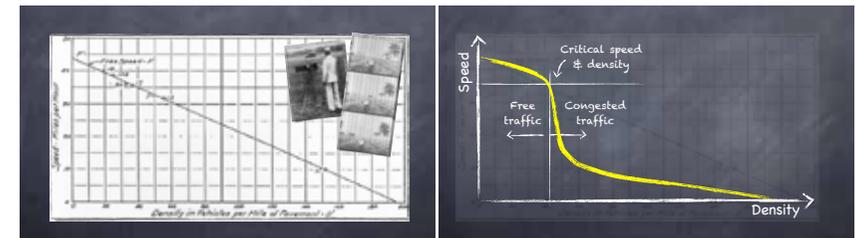


Figure 2

Over the years, the shape of this relationship has changed—new data has provided greater understanding. For example, we now know that during the transition from free-flowing traffic to congestion the speed drops very quickly (Figure 2 right) and that the form of the fundamental diagram determines the way in which perturbations propagate in traffic flows. The fundamental diagram is still one of the basic cornerstones of traffic flow theory. A second fundamental cornerstone (and the only real physical law) in our field is **the law of conservation of vehicles**. This law simply states that when more traffic flows into a specific road section than out of it, the number of vehicles on that section increases. In the 1960s, the mathematics that links this law to the fundamental diagram was developedⁱⁱ and with it the first-order traffic flow theory that we still apply today.

What the theory describes can be compared with what happens if we open a tap above an upturned bottle. If more traffic comes out of the tap than is able to flow through the neck of the bottle, the speed just above the bottleneck

ⁱⁱ *The mathematical tools have long existed and describe the dynamics of compressible fluids, but Lighthill & Witham (On Kinematic Waves II: A theory of traffic flow on long crowded roads. Proc. R. Soc A 229, 317 (1955)) and Richards (Shock waves on the highway. Operations Research 4, 42 (1956)) were the first to apply these to traffic.*

drops and traffic jams develop moving in the opposite direction to the flow. In this, the traffic flow is compressible—and here we can see Greenshields in action: the denser the flow (the shorter the distance gaps between vehicles), the slower they move and the redder the colour in Figure 3.

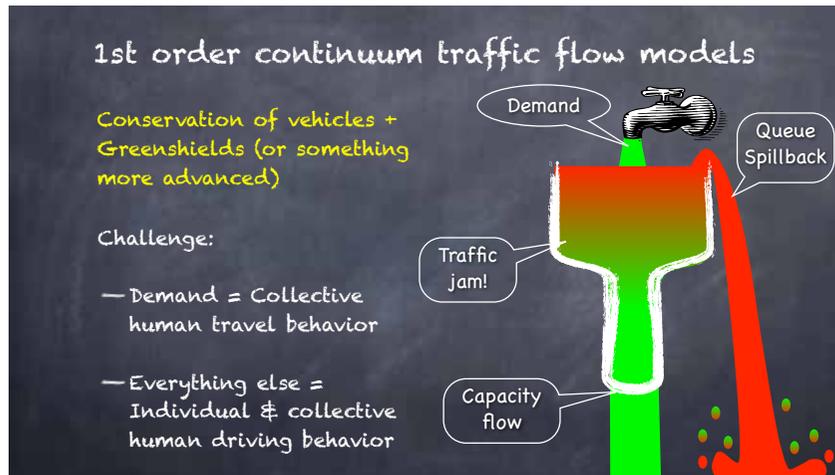


Figure 3

That what comes out of the tap is a result of collective behavior: where are we going, when, how and with what form of transport? What happens in the bottle itself is also the result of collective human (driving) behaviour. This means that the bottleneck capacity is not a characteristic of the road, but the result of driving behaviour: how close dare you drive behind the vehicle in front of you at high speed? In real traffic, another strange thing happens: as soon as a traffic jam develops, the bottleneck becomes slightly smaller: the so-called capacity drop. And there are many more obvious and not-so-obvious differences between the flow of traffic and fluidsⁱⁱⁱ. As a result, traffic tends to go from bad to worse: the greater the number of vehicles intent on passing through the bottleneck, the worse the end result is.

But more is happening than individual driving behaviour alone. There are also system dynamics: a traffic jam grows in the opposite direction of flow although the traffic continues to go forwards, of course. In networks, the way in which

ⁱⁱⁱ When you think about it, the continuum representation is of course nonsense. In a droplet of water, there are more water molecules than there have been or ever will be vehicles on the entire planet. The comparison between traffic and fluids therefore fails in all kinds of different ways. However, this approach can still be used to represent a number of phenomena effectively and we can also use it to make fairly sensible predictions.

perturbations (such as a traffic jam) move plays an extremely important role. Choice of route also plays a key role in networks: it determines how traffic flows spread across a network and in which bottlenecks problems could develop. You can therefore say that these system dynamics and these route-choice processes take place at a higher level of scale than driving behaviour.

MULTI-SCALE REPRESENTATIONS OF TRAFFIC AND NETWORKS

This brings us to an important point. Road traffic is a process that—like very many processes in the universe—can be described at many levels of scale^{iv}. The choice for a specific scale level depends on the phenomena one wishes to replicate. In this context, levels of scale are determined by:

1. How is traffic represented? Are we looking at the flow or at individual vehicles, and how do we describe the driving and travel behavior: individually or on average?
2. How is the infrastructure represented? Is it a one-to-one representation of the asphalt and the lines on it or do we use circles and arrows to represent the infrastructure, and if so what do they mean?

The first-order bottle approach just described is an example of a macroscopic traffic model, in which traffic is represented as a continuous (compressible) flow. If we wish to apply this kind of model in a network, we represent that network as a system of bottles and pipes linked together, through which the traffic literally flows, see Figure 4.

In order to simulate and predict traffic with this kind of model, we need data in order to ensure that the parameters in that model have plausible values. These parameters describe average driving behavior—these determine the shape of the fundamental diagram introduced earlier: capacity, critical speed etcetera. These kinds of parameters can be estimated relatively effectively using the data we collect, for example, from vehicle-detection loops and all kinds of other measurement systems. But we need much more data than that. How large is the traffic demand? Where is the traffic heading and via which routes? These kinds of things are much more difficult to estimate from data.

^{iv} These descriptions are typically not consistent, this is one of the focal points in the PhD thesis work of Mahtab Joueiai, see for example Joueiai, M., Van Lint, H., Hoogendoorn, S. Generic solutions for consistency problems in multi-scale traffic flow models - Analysis and preliminary results (2013) IEEE Conference on Intelligent Transportation Systems.

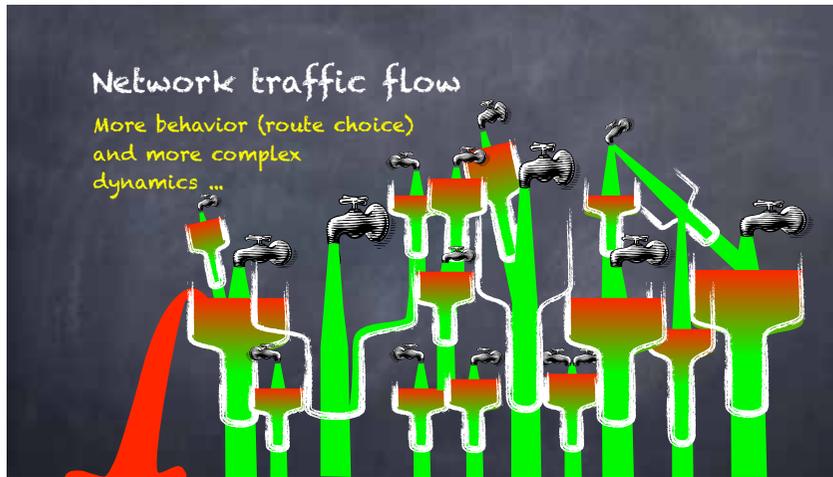


Figure 4

Consider this example. The traffic network in Figure 5 contains 3,250 origin-destination (OD) pairs, in other words combinations of places where traffic is coming from and going to. Now, we would like to know how much traffic there is for each OD pair, preferably also at different times of the day. Can we estimate these flows using data collected at the 350 locations where we actually count the vehicles? This is "difficult" without making an awful lot of assumptions^v.

At the end of 2014, one of my PhD students has been awarded with a doctorate for an approach that makes estimating and predicting traffic demand much easier^{vi}. She achieves this by using clever techniques to dramatically reduce the number of origin-destination relationships, sometimes by more than 90%. Using the remaining 50 relevant OD pairs, it becomes possible to predict the variation in overall traffic demand across the day. This then makes it possible to solve the problem, although it still involves the need to make a fair share of assumptions.

This approach of dramatically simplifying the problem because it would otherwise be insoluble or produce answers that have no predictive value, is an approach that can be used in a much wider sense. For example, if the precise dynamics of

^v This is a euphemism. The fact that there are more degrees of freedom than available data implies that the system is underdetermined. There is an infinite number of OD matrices (with 3250 cells) that could explain these same 350 measurements. The only way of actually selecting an OD matrix is, for example, by requiring that the OD matrix must be very similar to an historic OD matrix or by using additional data.

^{vi} Tamara Djukic has defended her thesis on online OD demand estimation and prediction November 18, 2014 (available via the TU Delft repository at <http://library.tudelft.nl>)

traffic are not very relevant for a specific application, it makes sense to describe traffic in networks much more coarsely than with first-order traffic flow theory.

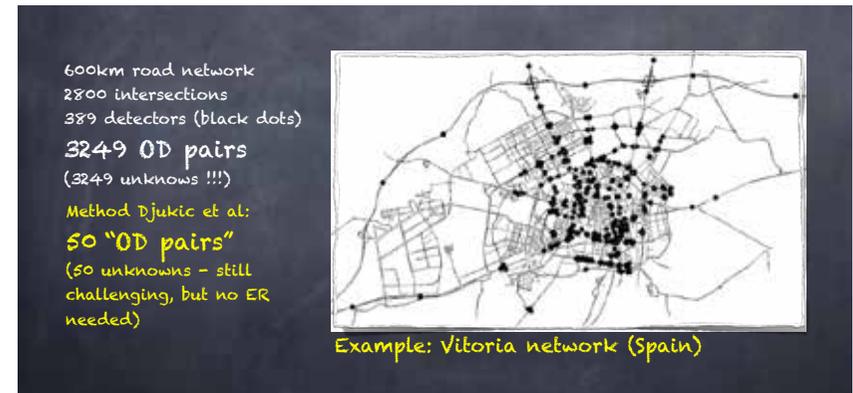


Figure 5

I call this approximate type of traffic model a **metascopic** traffic model (an "abstract" way of describing traffic operation in networks - see Figure 6). The idea behind it is that traffic operations in parts of a city or even whole cities can be modelled as a bathtub with taps and drains. What is remarkable is that at this level of abstraction, the description of network traffic operations turns out fairly simple. Academics, including colleagues here at TU Delft, have demonstrated that vehicle outflow from "the bathtub" grows with increasing vehicle accumulation, but that, beyond a critical vehicle accumulation, the outflow reduces to the point of complete gridlock^{vii}. Here again, we see the difference between traffic and water: also in whole networks traffic goes from bad to worse.

These bathtubs can be connected to one another using pipes through which the traffic flows. The result is a very coarse traffic model which can simulate major networks. A great deal more research needs to be done and much more data is required for it, but my colleagues are currently building this kind of model for the region of The Hague.

^{vii} See for example Geroliminis, N. and C.F. Daganzo, Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 2008. 42(9): p. 759-770; of Knoop, V.L., S.P. Hoogendoorn, and J.W.C. Van Lint, Routing Strategies Based on Macroscopic Fundamental Diagram. *Transportation Research Record*, 2012(2315): p. 1-10

Of course, one can also describe traffic in a much more detailed way, using what is known as a **microscopic** traffic model. In this kind of microscopic simulation, each vehicle moves according to its own behavioral models translated into mathematical rules. In a microscopic traffic model, the traffic network is also modelled in great detail because as well as going straight on, vehicles also need to be able to change lanes, drive across complex junctions, etcetera.

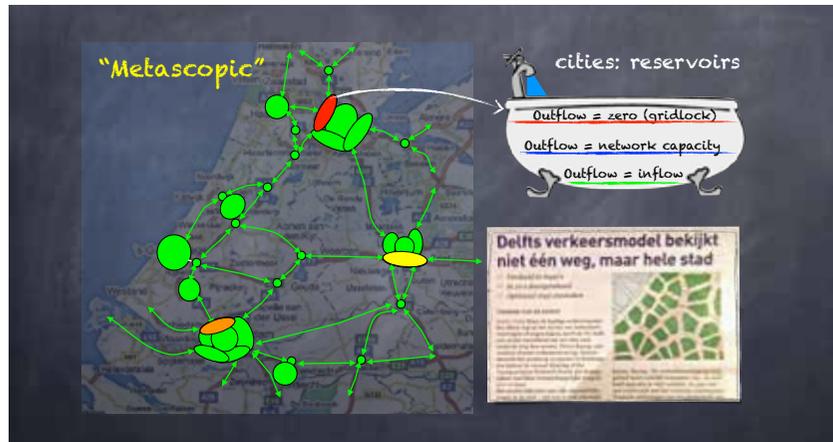


Figure 6

This of course provides a much richer description of traffic than a meta- or macroscopic description. If we then animate it in 3D, it even looks a lot like real traffic! But this detailed description comes at a high methodological price. Detailed models also include many more parameters (more degrees of freedom), all of which need to be estimated based on a lot of very detailed data, most of which we simply do not have—or at least not yet. This brings us to an important scientific principle that in practice is often mistaken for a paradox:

"more detail and more potential accuracy" does not necessarily result in "improved ability to predict".

On the contrary, simpler models (an upturned bottle) in many cases make better (more valid) predictions than more complicated (microscopic) models, because the former require many fewer assumptions. In other words, a greater level of detail primarily provides more opportunities for becoming lost in the fog!

SO WHERE DO WE STAND IN OUR FIELD TODAY?

Some good news first: there is considerable consensus about macroscopic models that describe equilibrium car-following behavior in all kinds of conditions. That may come as no surprise in view of the long history in that specialist area since Greenshields. All of these are descriptive theories that can mimic traffic (in a similar way to the upturned bottle), but provide no explanation for the underlying behaviour.

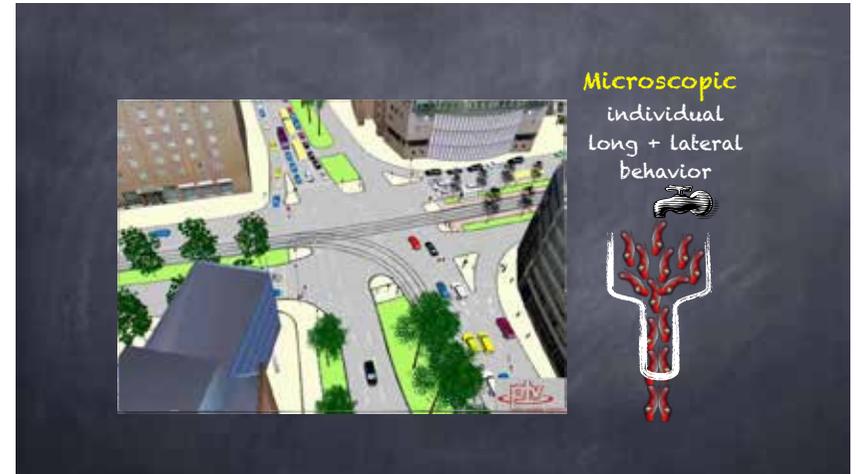


Figure 7 (source: <http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/use-cases/multimodal-systems/>)

There is less agreement about average "non-equilibrium behavior". Think, for example, of how traffic drives in and out of traffic jams and how perturbations emerge in traffic flows. The same applies to the dynamics at the network ("metascopic") level: this is a very lively area with a great deal happening at the moment.

However, there is no agreement whatsoever on the underlying individual driving behaviour. Some of the theories are descriptive; they can simulate driving behavior, but do not provide an explanation of why people drive as they do. Others contain tentative explanations, for example, how politeness plays a role in overtaking or how you work out in your head when exactly to cross an intersection. But the fundamental problem with all of these theories is that we simply do not have the data available to test whether these hypotheses about mutual behaviour actually hold water.

No misunderstandings here: micro-simulation provides fantastic opportunities for researching a large number of transport-related issues, which I will come to later, but if we want these simulations to predict real behavior, we first need to collect microscopic data on a much larger scale. Obviously, the next question is, how will we do that?

Broadly speaking, there are three ways in which we can collect data and I will present them here along two axes (Figure 8). The vertical axis tells us if we are actually measuring behaviour with the data and the horizontal axis tells us the extent to which we are aware of the conditions in which the data is collected. That is important because only then will we be able to repeat experiments and test whether the results really are significant.



Figure 8

The first option is to collect field data, as Greenshields did. This has been happening at TU Delft (and many other places) for a long time now: for example, we collect data from all the vehicle-detection loops on the Dutch motorways on the so-called regiolab-delft server in our department since 2000. Additionally we have a great deal more very detailed data. Of course, these data represent real behaviour, but it is very difficult to replicate the experimental conditions under which they were collected. For example, drivers do not have a sign on the roof saying where they are going, how long they have had their driver's licence or why they are driving as they are.

The second option is to conduct experiments in a laboratory. This is highly effective with pedestrians. Indeed, my colleagues have achieved worldwide acclaim in this area during the last decade. But putting city or motorway traffic into a laboratory is more difficult and I probably do not need to explain that any further.

This leaves option three: experimenting in a virtual environment, such as a computer, a driving simulator, a game. This provides complete experimental control. Driving simulators can for example be used to investigate how drivers respond to new infrastructure or how they deal with new gadgets on their dashboard. Multi-player computer games can be used to conduct research into the way in which travellers make decisions. The photographs in the Figure 9 are taken in Washington DC in 2012, where we conducted the Everscape experiment for the first time. In Everscape, 30 players are taken together to an island that will be hit by an earthquake followed by a tsunami five minutes later. Without the participants being aware of this, we investigated the extent to which the decision to run away is influenced by the bandwagon effect: will you run straight after the avatars in your immediate vicinity, or will you wait for more information, for example? We have now completed a whole series of experiments and the result will certainly make a great dissertation^{viii}!

The main question with these virtual experiments is of course which behaviour we are actually observing and how it relates to real behaviour. The answer to this is extremely nuanced... In certain cases, the perceived behaviour in a virtual world is only similar to real behaviour in terms of "direction", for example acceleration or deceleration, but not in terms of the precise amount. This varies for each type of behaviour and is also very dependent on the experimental design. The validity of driving simulator research for example is limited—for now at least.

What I would like to achieve is the same as all scientists who show a graph with two axes (Figure 8). I would like to reach the top right of the graph! Observe real behaviour with maximum experimental control!

In my view, we can get close to that in two self-reinforcing ways. First of all, by fusing field data obtained from all kinds of different sources. Think of data from vehicle-detection loop, satellite navigation systems, mobile telephones, social media, vehicle chargers, the weather, major events, etcetera. If we cleverly fuse

^{viii} For those who cannot wait: Mignon van den Berg, Eurico Doirado, Rob van Nes, Hans van Lint, Helmut Prendinger, and Serge Hoogendoorn. Application of a 3D multi-user environment for research on travel choice behavior. The case of a tsunami. 13th Int'l Conf of the Int'l Association for Travel Behaviour Research (IATBR'12), Toronto, Canada, 2012.7

these data, the resulting information represents much more than the sum of the parts. Let's call it BIG Data fusion. Fusing these different data provides the context that is needed in order to draw real conclusions about the underlying driving and travel behaviour and these are the two crucial ingredients to be able to simulate traffic on a large scale.

But even with BIG Data, it is not possible to measure everything. For example, you cannot measure how someone will respond to situations that (fortunately) happen only very rarely (accidents, for example) or situations that have never happened before. For that, you really need experiments in driving simulators (or games like Everscape). Look at this example: how will your driving behaviour change if 50% of all of the vehicles around you are driven automatically while the drivers calmly read a newspaper?



Figure 9

In 2013, the Dutch minister for Infrastructure and the Environment showed her guts by taking the passenger seat in a self-driving vehicle through normal traffic on the A10 motorway as part of the DAVI experiment (Figure 10). Fortunately, the experiment was coordinated by TU Delft and TNO, among others. However, three self-driving vehicles do not yet bring us into the future. I would for example very much like to be able to simulate and evaluate what the effects of an increasing number of intelligent vehicles on our road network would be. Will it be more efficient? Will it be safer? Where will the cut-off point be, at 10% self-driving vehicles? Or 50%? Before we can use a simulation model to calculate that, we first need to understand how drivers behave without that technology—

that will probably be very different from the situation now! And there are many more fundamental questions about the underlying driving behaviour that can only be investigated in a driving simulator.

Minister Schultz in zelfrijdende auto



Figure 10

That, in turn, means that the validity of these virtual environments needs to be significantly improved. This is an area with numerous opportunities that I would like to explore in the years ahead. Firstly, by conducting experiments in which many people participate simultaneously (a multi-player game approach): only then are we observing genuine interaction between traffic participants. Secondly, there are possibilities to make the virtual infrastructure and the behaviour of background traffic more realistic. One option that we have investigated recently is to have the background traffic learn a little from the drivers in the driving simulators.

RESEARCH PLAYGROUND: DELFT INTEGRATED TRAFFIC & TRAVEL LABORATORY

I would now like to return to the idea of a **playground**, in which researchers and students from lots of different backgrounds and disciplines can get their teeth into issues relating to traffic and transport. This is the kind of playground I want to build in the years ahead. The idea is to build a shared research environment in which data and simulation models are linked to each other. The glue between the data and the simulation models is a geographic description of the infrastructure and the built environment which can be easily translated into networks at different levels of scale. The key word is "OPEN", as evident in two areas:

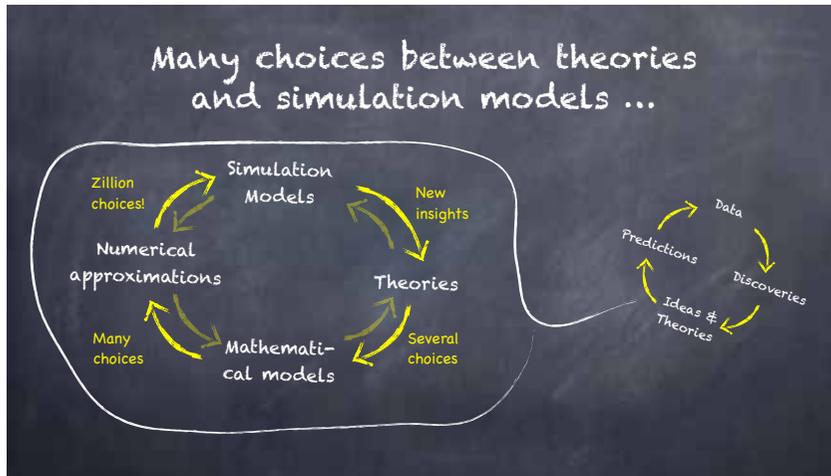


Figure 11

Firstly, I would like as much **open data** as possible in the playground and open standards for associating the data with networks and models. These standards really are crucial. They ensure that researchers can focus on the research rather than converting from data format A to format B. And it is not only researchers who benefit from interchangeable data and networks. Consultancies and their clients - governments - could find a better use for their time and money than wasting it on converting data and network descriptions from format A to format B.

Secondly, we want to further develop the OpenTraffic^{ix} **open-source multi-scale and multi-modal simulation model** initiative. Open source means that every researcher/programmer in our specialist area can contribute ideas and help to build the simulation models.

And that is very important.

This is because there are numerous steps between a theoretical idea and an actual simulation model (Figure 11). First, theories need to be translated into mathematics. These mathematics must then be solved numerically so that the model can ultimately be coded into software that can be run on a computer. Each of these steps involves numerous choices, which also means there are plenty of opportunities for making mistakes. It is therefore extremely important to be able to go back a step, to learn and apply new understanding and techniques.

^{ix} See <http://www.opentrafficsim.org>

Open-source development ensures that science continues to contribute to the development of simulation tools. There are already several successful open-source initiatives in our specialist area and I would like to make effective agreements with these pioneers on the input and output formats for their simulation models. I also hope to reach similar agreements with commercial simulation model builders.

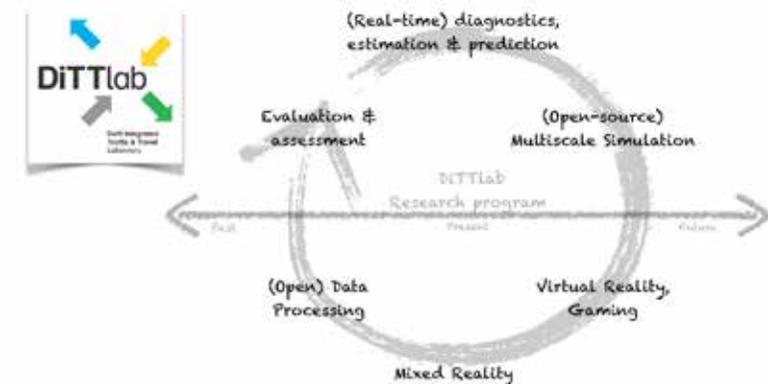


Figure 12. The two arrows pointing outward in the DiTTLab logo refer to what we want to get out of this playground: knowledge (TU blue) and applications with which more efficient, reliable and greener mobility becomes possible. The inward pointing arrows represent data streams from all conceivable sources. From januari 2015 onwards: www.dittlab.org

The research program associated with my AvL chair is called Delft integrated Traffic & Travel Laboratory (DiTTLab) and Figure 11 shows the main research themes of this program

I ordered the research theme's along a time axis from the past to the future. Field data obviously provide information about the past, given they are properly processed and fused using advanced techniques and tools. Simulation models on the other hand provide the means to predict future traffic operations, and can be used for example in "what-if" scenario analyses in which different possible futures are compared. I hope to have convinced you of the limitations of that approach due to the limited validity of the assumptions in these models and the data we feed into them. The only way to actually measure future behavior is virtual experimentation with driving and travel simulators. Clearly, also here, many puzzles related to the validity of those virtual environments need to be solved. By combining data and (simulation) models we can also reconstruct the traffic state at this moment, estimate the underlying causes for bottlenecks and make short term predictions of traffic and mobility patterns. One step further is

to use that estimated reality in a virtual environment in which researchers along with traffic managers and transport planners in real-time can compute and test new traffic measures. In the last part of this booklet I want to highlight two research themes sketched in Figure 11 (the blackboard on the right)



Figure 13

First I would like to move to **open-source multi-scale simulation models**. I have argued that we must choose the simulation model that matches the type of phenomena that we want to and can predict, based on the available data and the validity of our models. But one can also literally link models at different scale levels to each other. The result is something like the structure in Figure 13.

An important scientific puzzle in this kind of multi-scale model is how one can ensure that the models are consistent in relation to each other, such that that information flows effectively from one model to another. In a sense we are aiming to prevent a leak developing in the connections between the two different models. One of my doctoral candidates has already devised and tested a number of methods for this. For example, her research shows that the connection from one model to the other does not necessarily have to be in a fixed position. By allowing the connection to move depending on traffic conditions, you can at least ensure conservation of vehicles (one of the few things of which we really are certain).

This flexibility also makes it possible to switch dynamically and possibly also automatically between different levels of detail based on the phenomena

actually happening. This could be a solution to an even more complex puzzle: is it possible to automatically select the right scale level based on what happens in the network?

The number of potential applications is enormous. For example, we can create very coarse models of major cities or parts of cities using a metascopic model. The peripheral roads around them heading in both directions can be modelled macroscopically or microscopically, depending on the application. This then makes it possible to calculate the effects on such a network, for example, of a widened ring-road without needing to include a whole city in a micro (or macro) model. We select the level of abstraction based on the phenomena that we wish to make visible. Macro or Meta where possible and Micro where necessary.

We can learn a great deal from other disciplines that apply multi-scale approaches, such as climate modelling, material and structural mechanics and coastal morphology. The result is an integrated approach that makes the large-scale simulation of complex systems possible.

Finally, the research theme that has brought me the greatest pleasure in the last ten years and for which I have developed a lot of knowledge and practical tools together with doctoral candidates and colleagues: real-time traffic state estimation and prediction*. Some of these tools are already operating in practice, as part of the pilot project PPA (praktijkproef verkeersmanagement in Amsterdam). But there is much more in the pipeline and I am committed to make sure this knowledge will also find application in the traffic control centres in the Netherlands.

Figure 14 sketches a real-time traffic observatory in which many of these tools are used. The idea again is a playground, in which a simulated world is continuously reconstructed from real data, measured using all kinds of sensors. In order to make sense of these data, many tools are needed. The building blocks for all of these tools are traffic theory plus a hefty dose of statistics and mathematics. First of all to estimate the prevailing traffic conditions—especially in places where we do not take measurements. Secondly to make a diagnosis; is there for example a capacity problem on a particular route? Next we can use simulation models to do short term predictions and “what-if” analyses: what if we took this or that measure? What if an accident happens on a particular

* The following PhD students have worked in the past ten years on various puzzles related to traffic state estimation and prediction: Huizhou Tu, Hao Liu, Chris van Hinsbergen, Qing Ou, Thomas Schreiter, Yufei Yuan, Femke van Wageningen-Kessels - their theses are available via <http://library.tudelft.nl>

location?

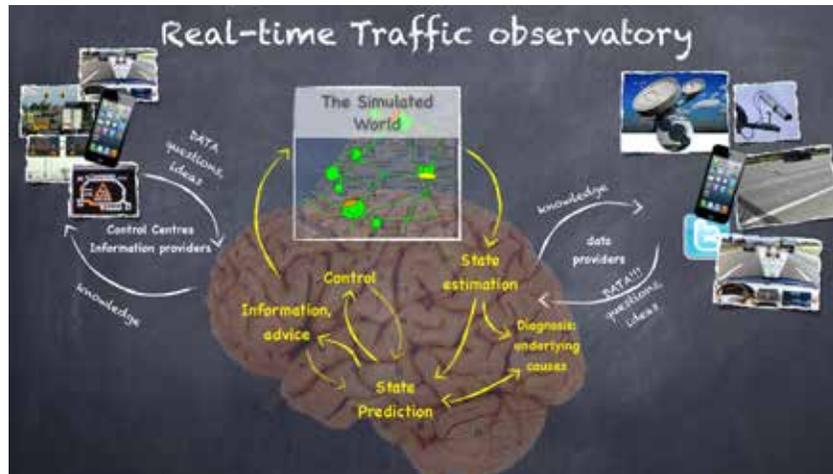


Figure 14

The Fastlane model I mentioned earlier was developed on my kitchen table in 2007 as part of a small project for the Directorate-General for Public Works and Water Management (Rijkswaterstaat). Fastlane then became the main subject in three doctoral theses, all of which were completed successfully in 2013. The knowledge acquired in these projects form foundations for such a traffic observatory, but many more puzzles need to be solved to actually build it. These Fastlane projects were implemented in collaboration with practice: the Rotterdam Port Authority, Rijkswaterstaat and the Rotterdam traffic company (a joint venture between the first two organisations and the City of Rotterdam). I hope to discuss and collaborate with these parties in projects to realize such a traffic observatory.

COOPERATION AND EDUCATION

To make a success of a research playground like DiTTlab, knowledge of traffic and transport is nowhere near sufficient. Almost all of the research themes are highly multidisciplinary in character. This means that there is a real need for multidisciplinary cooperation. This is a challenge that I very much relish.

The cooperation begins within our own department (Transport & Planning). From planning to operations and everything in between, it starts and ends with data. For a few years now, cooperation in the area of transport has been well

organised across TU Delft. Within the Transport Institute, I am now working with the Systems and Simulation lab at TPM to build the first OpenTraffic prototypes. A new development is the Institute for Advanced Metropolitan Solutions (AMS), as part of which TU Delft collaborates with the City of Amsterdam, MIT, Wageningen and many other partners. Within AMS, making sense of huge amounts of data will play a central role. I have the pleasure to direct with a few colleagues one of the first three projects within AMS, the Urban Mobility Lab, in which also data from all other (metropolitan) are fused and made available. The connection to DiTTlab is self-evident.

Also collaboration with the industry is important to me. Last year, I came into contact with CGI Nederland. After talking about data and simulation models for an hour or so, it was clear that collaboration would be to our mutual benefit. I am very happy with this collaboration that is based on a long term vision in which the different objectives of industry and academia are understood and respected. There are not very many private parties in our field in the Netherlands willing to make structural investments in innovation and cooperation with universities. I hope that DiTTlab will be a format that helps make this easier. This is essential, since long-term support is crucial for the survival of universities. Vice versa, structural investment in research and education offer huge potential returns.

Finally, if anything proves that multidisciplinary cooperation is fun and delivers results, then the inter-faculty programme TIL (transport, infrastructure and logistics) is it. This not only applies to the students but also the lecturers and other staff. What I am looking forward to most of all is making DiTTlab a playground for students. I hope that the data and models in DiTTlab will soon be made full use of in projects, courses and thesis subjects.

THANK YOU

That brings me to the end of my address ... Or at least almost, because I would like to offer my sincere thanks to a number of people.

I would like to start by thanking my closest colleagues within my research group and all the other colleagues in the Transport & Planning department, with whom I have worked with great pleasure for the last ten years and hope to continue to do so. The same applies to my colleagues in our CEG Faculty, at the Transport Institute and the TRAIL research school. I would also like to thank my colleagues and education support with whom I have worked in recent years with great pleasure as Director of Studies for the Master's degree programme in Transport Infrastructure and Logistics. And I would like to thank all the TIL

students who have shown trust in us and encouraged us to continue to improve the programme.



If anyone had told me in 1997 after getting my MSc degree that someday I would be making an inaugural address at TU Delft, I would have dismissed them as a nutcase. Applying the motto “science first, funding later”, my promotor Henk van Zuylen gave me 3½ year later an opportunity to work out what I actually wanted to do with my life. Thank you for that, Henk. It was the best career move ever. Many thanks also go to my other mentor Serge Hoogendoorn for his inspiration and friendship. A great many thanks to my doctoral candidates and students—thanks to you, I have one of the greatest jobs there could be. Several fantastic things happened in 2013. But it was also a very difficult year because I found myself hitting barriers that I could no longer think or regulate out of existence—my usual and so far successful strategy. I would like to thank my closest friends, my dear in-laws, my sister and family, and my parents for their unconditional support and love.

My father stopped working at the shipbuilding laboratory at TU Delft in 1986. One of the reasons for this were these “damned computers”; “Students will always eventually come to me if one of them malfunctions” and I can see it all too clearly. Dad had an infallible indexing system in his handwritten archive, behind and under which the contours of a regular office were still discernible. Dad, on behalf of the Van Lint family, I am now taking revenge by starting a

computer lab in which hopefully not nearly as many computers will malfunction. Finally, I bow in subordination to my greatest love, Christine. Although the chaos that I bring into your life would be too much for many to bear, you stand by me. Thank you for that.

I have said my piece.

**Faculteit Civiele Techniek
en Geowetenschappen**

Stevinweg 1
2628 CN Delft

Tel: +31 (0)15 278 9341

www.tudelft.nl