

# Taking deep uncertainty into account in traffic models

A case study of Groningen

I.H. van Baarle





TAKING DEEP UNCERTAINTY INTO ACCOUNT IN TRAFFIC MODELS:  
A CASE STUDY OF GRONINGEN

A thesis submitted to the Delft University of Technology in partial fulfillment  
of the requirements for the degree of  
Engineering and Policy Analysis Master of Science in Engineering and Policy Analysis

by

Ilmo van Baarle

August 2021

Ilmo van Baarle: *Taking deep uncertainty into account in traffic models: a case study of Groningen* (2021)

© ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

ISBN 999-99-9999-999-9

The work in this thesis was made in the:

Faculty of Technology Policy and Management  
Delft University of Technology

Supervisors:	Dr.ir. J. H. Kwakkel Dr. Y. Huang
Chair:	Prof.dr.ir A. Verbraeck
Company supervisor:	Ir. S. Chowdhury

# EXECUTIVE SUMMARY

Excessive car use is one of the main factors of congestion, global warming and bad air quality in big cities. Therefore, many municipalities are trying to ban out cars from their cities. To see which policies can achieve this change in traffic, complex macro-level traffic models that calculate congestion and modal split are created. More often than not, these models make predictions based on a small set of scenarios. However, in the rapidly changing times of today, these scenarios are often not sufficient enough to account for the highly uncertain future. Therefore, in this thesis, a traffic model from a case in the municipality of Groningen is tested on the unpredictable future more broadly. This is called a deep uncertainty analysis. The main research question is:

*“How to account for deep uncertainty in macro-level traffic models to achieve more robust decision alternatives?”*

In order to answer this question a fitting method to apply to the case is found. After that the model is defined, with the important uncertainties, policies and outcomes. Then the results are analysed to end with a discussion and conclusion of these results.

The futures in model studies are described by uncertainties. Uncertainties are defined in 5 levels of which level 4 and 5 are deep uncertainties. The pandemic that Covid-19 caused is a good example of such a high level uncertainty. A prominent method to take these uncertainties into account is Robust Decision Making (RDM). This is a 5-step model cycle in which the first step consists of defining the model. The second step runs this model over a wide set of futures described by the possible uncertainty values of the model. In this thesis, these futures are drawn via Latin Hypercube Sampling. In the third step, the relative effect of all uncertainties of all outcomes is calculated. Additionally, scenario discovery techniques can show which areas of futures is causing the system to fail or succeed. The fourth step is called the trade-off analysis. Based on robustness metrics, different policies are compared on performance. The robustness metrics are chosen based on the level of risk aversion of the decision maker and the preference towards a policy that suffices a threshold or a policy that performs well compared to other policies. In the last step new futures and policies are created. This research mainly focuses on the first four steps.

The idea of RDM is that a very broad set of scenarios is run. However, macro-level traffic models are often very detailed and complex which causes run time to be long. Take for example the model of the case study, which runs in 22 hours and 45 minutes. This is too long for the time in which an assignment must be complete and therefore the run time is reduced. The main place in which time can be reduced is in the procedure sequence. There are many procedures that are not important to the goal of this study. Due to the previous mentioned characteristic of macro-level traffic models, reducing the model detail is a more challenging task. The high amount of underlying relations makes changes very unpredictable. Eventually the run time of the case has been brought back to 12 minutes. The software in which the case runs, PTV Visum, also allows for 5 simultaneous runs. This means that the total run time of the broad set of scenarios is reduced by a factor 5. The next obstacle is the fact that the case runs in PTV Visum and the main method to perform RDM, the EMA.workbench, runs in Python. Via a COM interface these two models are connected.

In the first step of the RDM, the model is defined by a set of uncertainties, policies and outcomes. In total 5 uncertainties, 4 policies and 9 different outcomes are defined. The most important outcome is the total distance travelled by car. This is an outcome which is often used in traffic research. However, the policies defined are all fixed. Therefore, a set of changing policies is also defined. The changing policies allows the researcher to get a more clear view on correlations between policies and outcomes, but also between policies and uncertainties. The changing policies are the same as the fixed policies, beside the fact that they can change in value. In the second step of the RDM cycle the model is run over a wide set of plausible futures. In this case 1000 runs are performed. Every run describes a different future based on a combination of uncertainty values between a lower and upper bound.

The analysis of the data with feature scoring, PRIM and 4 different robustness metrics, showed the value of deep uncertainty analysis for macro-level traffic models. It can find scenario regions of interest, show interactions between uncertainties and outcomes and the policies can be ranked based on robustness. However, it also showed some characteristics of these types of models that limit the usefulness. The first limitation is the fact that most macro-level traffic models are over fitted. This causes the effect of a change in a single variable to be negligible. Over fitting is most obvious in the feature scoring, where policies have very little interaction effects with the uncertainties. This causes, no matter what the scenario is, the ranking of the policies to be exactly the same. The second limitation is the fact that the set of variables is limited. Something like costs is not included in the model and due to the high complexity it is hard to add this to the model. This limits i.e. a trade-off analysis, since the modeller is only possible to see trade-offs, based on different robustness metrics, within a single outcome of interest and not between the robustness of conflicting outcomes of interest.

There are many different ways future work could be built upon this work. The first is to apply the RDM cycle to a macro-level traffic model with conflicting outcomes of interest. The second is to perform an optimization on changing policies with more advanced techniques, like Multi Objective Robust Decision Making. The biggest obstacle to apply this is the total run time of macro-level traffic models. Therefore, computational power has to be increased or computational expensiveness has to be decreased. The first can be achieved by, for example, using nodes from providers like Amazon web services. The most plausible way to achieve the latter is by completely remaking a model from scratch. The third option is to build a less complex model which represents the behaviour of a big macro-level traffic model. This introduces the last possibility for future work, multi-resolution modelling. By making both a simple small and a big complex model, policies can be tested on a broad set of futures in a small model, after which a big model runs these policies for more in depth results. However, the question arises whether or not such a big model is needed. Due to the high level of uncertainty in these models, very exact calculations are still uncertain. A in-depth deep uncertainty analysis might prove more useful than a small set of scenarios in a complex model.

## ACKNOWLEDGEMENTS

During my bachelor "Technische Bestuurskunde" I started to develop a keen interest in complex transport systems and modelling and simulation. The master engineering and policy analysis allowed me to pursue these two passions and made me better and more knowledgeable in these subjects. One of the most interesting courses of my master period was "model based decision making", a course taught by Jan Kwakkel. This thesis applies the methods learned in that course.

The last couple of months were a little different than I would have expected writing a thesis to be. Almost all work has been done from home. However, with the close help of some people I have been able to finish in time. First of all I want to thank Jan Kwakkel for the inspiration and help during our meetings. Thereby, Jan has had a big contribution in writing the connector between Visum and Python. Even though meeting in person was not a option, the guidance felt really personal. I also would like to thank Alexander Verbreack and Yilin Huang for their work as chair and as second supervisor respectively.

This research was conducted in cooperation with Sweco. The people at Sweco have done everything in their power to involve me during my internship. Especially Saidul who was my supervisor from Sweco. Our weekly Monday morning meeting have been a great start of the week. You have made sure I came in contact with the right people and tried to help me where it was needed. You always pushed me to follow my time schedule which is part of the reason I finished in time.

Lastly, due to Covid-19 it was hard to focus at home. Luckily (for me) my friends had the same problem. Around 3 times a week we all went to the VU Amsterdam to work on our colleges, projects and thesis. I am thankful for the mental support we could give each other.

I hope you enjoy reading my thesis.

Ilmo van Baarle





# CONTENTS

1	INTRODUCTION	1
1.1	The negative effects of car use	1
1.2	Traffic modelling	2
1.3	Deep uncertainty in transport modelling	3
1.4	Research scope	4
1.5	Research questions	4
1.6	Research approach	5
1.6.1	Mode choice	5
1.6.2	Route assignment	6
1.6.3	Run time	7
1.7	Structure	7
2	METHODOLOGY	9
2.1	Uncertainty is everywhere	9
2.2	Ways to deal with uncertainty	10
2.2.1	Robust decision making	11
2.3	Conclusion methodology	16
3	MODEL IMPLEMENTATION	19
3.1	Adapted model	19
3.1.1	Considerations to make when adapting the model	19
3.1.2	Changes in the model	19
3.1.3	Validation of the adapted model	21
3.2	Connection between Visum and the EMA_workbench	21
3.3	Step 1: Decision framing	23
3.3.1	Externalities	23
3.3.2	Levers	24
3.3.3	Changing levers	25
3.3.4	Outcomes	25
3.4	Conclusion model implementation	26
4	RESULTS	29
4.1	Step 2: Experimental design	29
4.2	Step 3: Global sensitivity analysis	29
4.2.1	Feature scoring	30
4.2.2	Step 3: Scenario discovery	33
4.3	Step 4: Robustness analysis	39
4.4	Conclusion	41
5	DISCUSSION	43
5.1	Interpretation	43
5.2	Implications	44
5.3	Limitations	45
5.4	Future work	47
6	CONCLUSION	49
6.1	Answer research question	49
6.2	Recommendations municipality of Groningen	51
A	POSSIBLE MEASURES	57
B	FEATURE SCORING	59
B.1	30 km policy	59
B.2	Bike policy	59
B.3	Parking policy	60
C	PRIM	63
C.1	Peeling and covering process	63



# LIST OF FIGURES

Figure 1.1	Groningen Visum model view . . . . .	6
Figure 2.1	Locations of uncertainties in policy analysis (Walker et al., 2013) . . . . .	10
Figure 2.2	5 steps of a RDM (Lempert et al., 2013) . . . . .	12
Figure 2.3	Example of a 2D LHS (van der Scheer, 2021) . . . . .	14
Figure 2.4	Robustness metrics as defined in McPhail et al. (2018). The bar goes from green to blue, where green is a low level of risk aversion and blue is a high level of risk aversion . . . . .	17
Figure 3.1	Comparison of the original and adapted model based on three different outcomes . . . . .	21
Figure 3.2	Conceptual connection between the EMA_workbench and Visum . . . . .	22
Figure 4.1	Feature scoring with policies . . . . .	30
Figure 4.2	Feature scoring in the base case . . . . .	31
Figure 4.3	Parcoord plots of the feature scoring of all uncertainties on three different outcomes . . . . .	32
Figure 4.4	Feature scoring with changing policies . . . . .	32
Figure 4.5	Bar chart with the 95% confidence intervals and the mean of three different sets of feature scores . . . . .	33
Figure 4.6	PRIM in the base case for the 10% worst performing scenarios	34
Figure 4.7	PRIM in the base case for the 10% best performing scenarios .	35
Figure 4.8	PRIM for the 10% worst performing scenarios with policies .	36
Figure 4.9	PRIM for the 10% best performing scenarios with policies . .	37
Figure 4.10	PRIM for the 15% worst performing scenarios with changing policies . . . . .	38
Figure 4.11	PRIM for the 10% best performing scenarios with changing policies . . . . .	40
Figure A.1	Possible policies . . . . .	57
Figure B.1	Feature scoring with the 30 km policy . . . . .	59
Figure B.2	95% confidence interval of the feature scoring of car_total for the 30 km policy . . . . .	59
Figure B.3	Feature scoring with the bike policy . . . . .	60
Figure B.4	95% confidence interval of the feature scoring of car_total for the bike policy . . . . .	60
Figure B.5	Feature scoring with the parking policy . . . . .	61
Figure B.6	95% confidence interval of the feature scoring of car_total for the parking policy . . . . .	61
Figure C.1	Conceptual illustration of PRIM's: a) peeling and b) covering process. (Lempert et al., 2008) . . . . .	63



## LIST OF TABLES

Table 1.1	Regions in the model . . . . .	6
Table 3.1	Steps taken to adapt model . . . . .	20
Table 3.2	Uncertainty ranges . . . . .	24
Table 3.3	Ranges of changing policies . . . . .	25
Table 3.4	Metrics regarded in the model . . . . .	26
Table 4.1	Robustness scores, between brackets show if a high or a low value is good . . . . .	39
Table 4.2	Feature scoring on regret . . . . .	39

# 1

## INTRODUCTION

### 1.1 THE NEGATIVE EFFECTS OF CAR USE

Cars are one of the most used vehicles for transportation. However, there are many negative effects regarding car use. The main negative externalities are the emission of CO<sub>2</sub>, NO<sub>x</sub> and PM<sub>10</sub>. This causes global warming and harm to public health (McMichael et al., 2003). There are already many policies made to cope with these negative effects. For example, the ban of cars with a euro 5 or lower emission standard in Berlin or the introduction of emissions free zones in many big cities (Sheahan, 2018). Beside the negative effects of singular car use, there are also some effects based on the amount of car use. Congestion is the main negative effect caused by this. Due to the capacity of roads being lower than the volume of vehicles driving on these roads there are 63.5 million hours of congestion in the Netherlands in 2020, which was a year with a relatively low amount of congestion (Rijksoverheid, 2020). Next to the congestion there are parking problems in large cities, but smaller towns as well (Ibrahim, 2017).

The singular car use problem mainly causes negative effects on health and global warming (Douglas et al., 2011). This is most obvious in cities like Beijing where the large amount of emissions causes a thick layer of smog above the city (Peng et al., 2020). As a result of the smog many citizens wear face masks to prevent health issues and tourists think twice before traveling to Beijing. Global warming is one of the biggest challenges the world faces. It speaks for itself that the negative effects this can have are enormous. The capacity problem on both roads and parking space partly adds to the above painted problems, because of cars having to pull up and stop more often. It also has some second order negative effects. Due to congestion or people unable to find a parking spot, efficiency of traffic is lowered. Thereby, trucks, which are the main transporters of goods, get a lot of delays.

There are many uncertainties and different possible policies accompanied with this problem. It is unsure what the future in transportation brings (Lyons and Marsden, 2019). The upcoming electric vehicles, the evolution of Mobility as a Service (MaaS), hydrogen driven cars and the reductions in emissions for gasoline driven cars. Partly due to these new developments, the list of possible policies to reduce car use and emissions is almost endless. Different roads can be changed or laid down, a shared bike system can be introduced or the speed on a road can be changed. The environment in which these policies have effect can also change in ways that are hard to predicted. For example, a situation in which they lay down an extra lane for a highway. On the one hand, the impact of this extra lane can be very big if the population around this highway increases or if this population gets a higher percentage of commuters. On the other hand, this impact can be negligible if the volume on the road does not change or even decreases. There are many external factors that determine these variables. A fixed number struggles to capture the complete picture. Therefore, a range of possible values covers the future better than a fixed number.

Even though the future is very uncertain, models that value these policies are not taking these uncertainties into account in a proper manner. In many studies a set of 3 to 10 different scenarios are researched. In very complex systems these scenarios only show a small part of the high amount of possible scenarios. Due to this, policies tested can have completely different outcomes than expected. A

policy that performs well in 3 different scenarios can perform quite poorly in 500 other scenarios. Therefore, taking these uncertainties into consideration in a broader manner can prove helpful in finding a more robust policy.

## 1.2 TRAFFIC MODELLING

There are multiple forms of traffic modelling and simulation. [Ratroun and Rahman \(2009\)](#) describe three different levels of traffic modelling: microscopic, mesoscopic and macroscopic. Microscopic modelling focuses on the individual behaviour of the entities in the model. Macroscopic modelling is based on the continuum traffic flow theory whose objective is mostly based on the speed, density and volume of the total traffic. This means that there is little attention to the interaction between agents. Lastly, mesoscopic modelling is a combination between microscopic and macroscopic. It tries to catch the dynamics of the micro model while focusing more on the continuum traffic flow. Sometimes transport models of different levels are combined. [Yang and Morgan \(2006\)](#) presented a model that was microscopic on the points of interest and meso- macroscopic at the surrounding areas that were of less interest. By doing this, a bigger network can be stimulated without losing too much computing power.

Traffic and transportation systems are complex systems, due to the fact that they are characterized by a great amount of factors and their relationships ([Barceló, 2010](#)). The complexity increases further by the fact that these factors and relations have a lot of uncertainty ([Ottomanelli and Wong, 2011](#)). This uncertainty stems from the fact that in these types of models human behaviour, environment and external effects play a big role. Take for example the effect of COVID-19 on a system, this is often not possible to predict perfectly. In order to have more reliable solutions, policymakers have to take this uncertainty into account.

[De Jong et al. \(2007\)](#) describe how uncertainty is taken into account in traffic forecasts. They consider two different forms of uncertainty, model uncertainty and input uncertainty. Model uncertainty is mostly based on the relationships within the system. In order to account for the input uncertainty they use Monte Carlo sampling. This method is put forward by multiple papers ([Hugoson et al., 2005](#)) ([Zhao and Kockelman, 2002](#)). For a part of the model uncertainty they also use Monte Carlo, but also for a part the bootstrap technique. The bootstrap technique picks, from a predefined  $x$  set of values,  $x$  new values where the same value can occur more than once ([Hugoson et al., 2005](#)). By applying these techniques, the forecast gets a wide range of possible values which represent the uncertain future. Another way of sampling the uncertainties is through Latin hypercube sampling (LHS). This technique tries to give an equal distribution of uncertainty values over scenarios ([Sheikholeslami and Razavi, 2017](#)).

[De Jong et al. \(2007\)](#) continue their paper with a case study of the Dutch national model system, called "Landelijk Model Systeem" (LMS). The LMS is a model that contains the whole Netherlands with more than 1300 zones and their OD matrices focusing on passenger transport. The model calculates growth factors of each origin-destination relation by mode, purpose and time. The growth factors are of a future year based on a base year. Uncertainty is tested for the tour frequency models and the mode destination models. The values of the uncertainties in the scenarios are picked by Monte-Carlo sampling. First a distribution of an uncertainty was calculated by checking data from the past 30 years. Based on this distribution values are picked. In total 100 scenarios were calculated of whom 50 contained a policy and 50 that contained no policy. The conclusion of the model was that input uncertainty had the most influence of the outcomes. This contained variables like "car ownership" and "income distribution". The outcomes were expressed in the standard distribution between runs.

### 1.3 DEEP UNCERTAINTY IN TRANSPORT MODELLING

There is a difference between stochastic uncertainty and deep uncertainty. [Lempert \(2003\)](#) describes this difference as follows: “The condition in which analysts do not know or the parties to a decision cannot agree upon (1) the appropriate models to describe interactions among a system’s variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes”. This is especially true in model-based decision support for complex systems ([Lempert, 2002](#)). [Walker et al. \(2012\)](#) describe 5 levels of uncertainty, with level 1 being the lowest level and level 5 the highest. A level 1 uncertainty can often be dealt with through a sensitivity analysis. A level 5 uncertainty is an uncertainty of which we know that we do not know it. Think about natural disasters which we did not expect. [Walker et al. \(2012\)](#) assign only the level 4 and level 5 uncertainties to deep uncertainties.

The notion of uncertainty is often mentioned, but not often dealt with extensively ([Rasouli and Timmermans, 2012](#)). Increasingly, studies acknowledge the importance of uncertainty of all the three levels of transport models. It is also mentioned that policy makers should not look at the most plausible future, but rather at a range of possible futures they can encounter ([Curtis et al., 2020](#)). [Petrik et al. \(2018\)](#) did an uncertainty analysis on an agent based model (micro simulation) of Singapore. A model with more than 600 uncertainties first got reduced to 100 based on a sensitivity analysis. This was possible with so many uncertainties due to the fact that the model was divided into multiple sub-models. This reduced the amount of simulations needed to be run and uncovered the most influential uncertainties. After that, based on latin hypercube sampling, 500 scenarios were run. The uncertainties used are mostly of level 1, 2 or 3. Therefore, it is not in essence a deep uncertainty analysis which causes some deep uncertainties to be not dealt with enough. Adding to this, running only 500 scenarios with a set of 100 uncertainties is very little. The amount of combinations possible are far greater. Therefore, it is even questionable if 10000 scenarios would suffice.

The way researchers quantify uncertainties in transport modelling differs between each other. Most of them quantifies it via link flows ([Lowe et al., 1982](#)). Another way of quantifying it is via travel times or emissions which need an extra factor, like emissions per kilometer, in order to quantify ([Schrijver et al., 2003](#)). A more aggregated way of quantifying the effect of uncertainty is by taking the amount of trips or passenger kilometers per mode ([Armoogum, 2003](#)). These quantifications are usually presented by giving the 95% with the variance and the standard distribution. This accomplishes a percentage based likelihood for a scenario.

[Calvert et al. \(2018\)](#) research the effect of uncertainties in traffic planning. They noticed that the average of scenario inputs are not the average of scenario outputs. This is caused by the fact that there are many secondary effects like congestion. This means that in a big part of the scenarios there is no heavy congestion, while congestion is one of the main reasons for traffic planning. Therefore, focusing on these scenarios can prove more useful than focusing on the scenarios in which traffic planning is not needed.

[Overtoom et al. \(2020\)](#) assessed the impact of shared autonomous vehicles on congestion and curb use. Again, they acknowledge the presence of deep uncertainties, however they do not assess these uncertainties thoroughly. This seems to be the case in most of the researches found. Deep uncertainties are acknowledged, but weakly accounted for by for example a sensitivity analysis. This makes the presented solutions less robust than could be possible.



## 1.4 RESEARCH SCOPE

Traffic models are usually representations of very complex systems, this is even more the case in macro-level traffic models. Macro-level traffic models are therefore dependent on deep uncertain factors. These uncertainties, which are mainly input uncertainties, are often not properly accounted for. Some researchers have tried to account for uncertainties in macro-level traffic models but have either failed in acknowledging the deep uncertain nature or have tried to account for it by running just a small set of scenarios. This research applies the Robust Decision Making (RDM) cycles in a case study to account for the deep uncertainties in a macro-level traffic model.

## 1.5 RESEARCH QUESTIONS

A lack of regarding deep uncertainty in traffic models makes the reviewing of policies a less robust process. Therefore, In this research, the first steps into regarding deep uncertainty in traffic models are taken. The research question that fits that purpose is:

*“How to account for deep uncertainty in macro-level traffic models to achieve more robust decision alternatives?”*

In order to answer the main research question, some sub-questions are defined. These sub-questions will divide the research in multiple segments which can jointly answer the main research question. Below every sub-question is listed and described. They will also be linked to each other to see how they can answer the main research question. Some sub-questions are based on the case study defined. The sub-questions go along the line of the Exploratory Modeling and Analysis (EMA) methodology defined by [Bankes \(1993\)](#).

- *Which method can be used to take deep uncertainty into account in a traffic model?*

This sub-question has little to do with the EMA methodology. The purpose is to extend the basis on which the research is conducted. There is a lot of theory about how numerous deep uncertainty methods can improve decision-making. Some of these methods are fitting for traffic models and some are not. The eventual outcome of this research will depend on these methods.

- *What are the important deep uncertainties, policies and outcomes in traffic models?*

Among the EMA methodology three steps are identifying the (deep)uncertainties, the design of policy levers and the identification of troublesome/promising regions. The uncertainties are set at the early stages of the modeling process while the outcomes and policy lever are able to change based on outcomes. As mentioned by [Walker et al. \(2012\)](#), only level 4 and level 5 uncertainties are regarded as deep uncertainties. Therefore, the uncertainties found have to be evaluated on their level. The policies that are possible to input in the model will be based on the resources of the problem owner of the case that is studied. Regarding the outcomes, not only the types of outcomes are important. For example the maximum acceptable values are of importance to find regions of interest.

- *Which uncertainties and policies is the model sensitive to?*

If all the uncertainties are implemented the model can produce outcomes. This gives a first insight into the uncertainty space and which regions produce interesting outcomes. This part of the EMA methodology is iterative. Policies will be adjusted

based on the interesting regions. The outcomes differ from the past results due to the changed policies. This shows new interesting regions on which again the policies can be adjusted.

- Which policies perform best across a wide set of future scenarios?

The result of the other sub-questions lead to a set of policies that perform best in certain situations. Based on the answer of sub-question 1, a conclusion can be drawn about which policy proves most useful for this case. The word “best” thus has to be defined by the results of sub-question 1.

## 1.6 RESEARCH APPROACH

The goal of the research is to analyse how deep uncertainty can be taken into account in a decision based on a traffic model. Yin (2017) emphasizes that in order to test if a certain method works, a case study is a good approach. Therefore, the research will be conducted on the basis of a case study. The case is based on a research currently conducted by Sweco for the municipality of Groningen. Groningen is developing the *mobilitieitsvisie* for 2030 and 2040. The *mobilitieitsvisie* is a report about the future of transportation in the municipality. It contains for example the percentage of car users per age and the kilometers traveled by bicycle. In cooperation with Sweco, a model in PTV Visum, which represents the traffic flows on a small scale within Groningen and on a big scale around Groningen, is created. The model allows different policies to be inputted. It can be classified as a macro-level model, due to the high level view on the road system. The main goal for the municipality of Groningen is to reduce the amount of kilometers travelled by car. These kilometers should be transferred to for example bicycle kilometers or public transport kilometers.

The Visum model that is used is designed by a collaboration between 4cast and Sweco for the municipality of Groningen. It contains the details of all the roads and all the demographic data of Groningen in detail. A model view is given in figure 1.1. Data of the rest of the Netherlands and foreign destinations is less detailed, but it is taken up in the model. The demographic data is described in more than 2200 zones defined in the model. The model is based on the four-step travel model (McNally, 2007). The four steps are trip generation, trip distribution, mode choice and route assignment. The trip generation part determines trips from origins to destination for each zone based on demographic and socioeconomic data. Trip distribution matches the origins with the destinations using a gravity model. This takes into account the travel costs between origins and destinations and the relative activity at them. The latter two parts are described in more detail in the next to sub chapters.

### 1.6.1 Mode choice

The mode choice part of the model determines the proportion that transport modes are used between origins and destinations. The modal model is of the multinomial logit form (Kwak and Clayton-Matthews, 2002). It uses the first two steps and data about transportation modes and their usage costs as input. In the Visum model of Groningen the mode choice is done for single movements and for chains of movements. A single movement is for example from home to the supermarket and back. A chain of movements is for example from home to the supermarket, from the supermarket to school and from school back home. These movements are calculated in 15 different procedures in the procedure sequence of the Groningen model. The first procedure, which calculates the mode choice for the non-chain movements, accounts for around 80% of all the movements, which means that around 20% are chain movements. The mode choice model results in two dataframes which contains

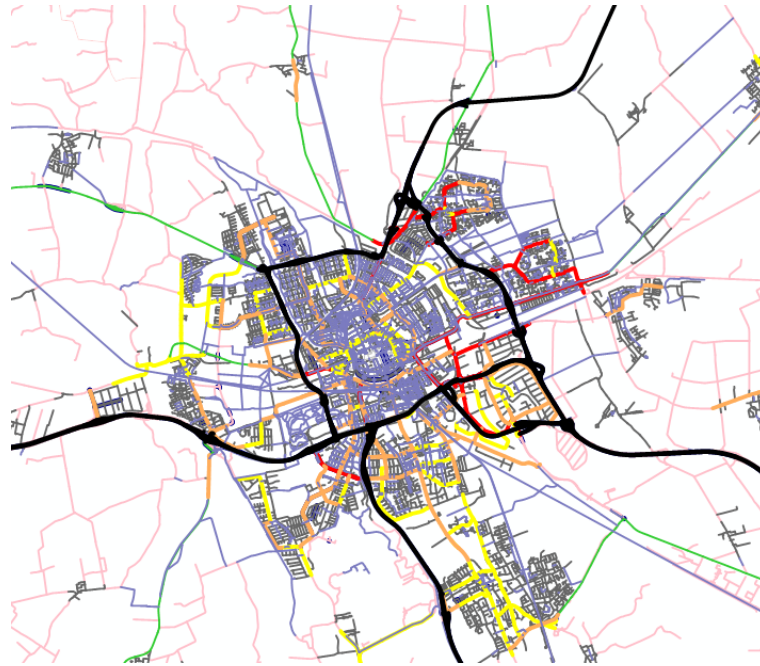


Figure 1.1: Groningen Visum model view

Table 1.1: Regions in the model

1	Municipality Groningen
2	South
3	Southeast
4	Northeast
5	Northwest
6	Southwest
7	Other parts of the Netherlands
8	Abroad
$\geq 2 \leq 6$	Region outside municipality
$\leq 6$	Region Groningen

the amount of movements per travel purpose per transportation mode from every region to every other region (table 1.1).

### 1.6.2 Route assignment

The route assignment part allocates a route for every origin destination pair combined with the mode choice. There are several ways this can be done. Two examples are the Wardrop equilibrium assignment and the ICA (Intersection Capacity Analysis) assignment. In the Wardrop equilibrium assignment every car chooses the shortest route, subject to the other drivers. The difficulty here is that drive time is influenced by demand and demand is influenced by drive time. This is called a bi-level problem. Therefore, there are multiple iteration needed to find a good balance. ICA assignment is a combination between an equilibrium assignment and node impedance calculation (PTV Visum 2021 - Manual). These two parts influence each other in an iterative manner. The assignment calculates the volume on node and with this information the delay time and capacities of nodes/turns is calculated. The ICA assignment is usually more demanding than the Wardrop equilibrium assignment.

In the Groningen model the route assignment gives a clear overview of the business of roads. For some parts it uses the ICA assignment and for some parts the

equilibrium assignment. It gives a map where the roads are colored based on the amount of cars or congestion on a certain street.

### 1.6.3 Run time

The Groningen model has a run time of around 22 hours. The first two parts of four step travel model take very little time. The mode choice model takes around 2 hours and the route assignment model around 5 hours. However, the last two parts are run 3 times to account for deviations in the results. These numbers are retrieved from an 80 core computer. For many methods that take deep uncertainty into account a lot of different runs have to be conducted. It is clear that the run time of the Visum model is too high and therefore is one of the key obstacles in this research.

## 1.7 STRUCTURE

The remainder of this thesis has the following structure. Chapter 2 describes the methods used to deal with deep uncertainty. Chapter 3 first addresses the run time of the model and then a connection between python and Visum is made. Finally, the method describe in chapter 2 is applied to the model. Chapter 4 discusses the results from the analysis. In chapter 5 the discussion of the results is given, which include the limitations of this work and ideas for future work. Lastly, a conclusion is given in chapter 6. This includes the answering of the research questions and a recommendation to the municipality of Groningen.



# 2 | METHODOLOGY

This chapter dives deeper into the notion of deep uncertainty and the value within policy making. A prominent method from the literature, called robust decision making, is discussed. Thereby, methods for global sensitivity analysis and scenario discovery are given. The chapter is finalized with a discussion of different robustness metrics, which are used to evaluate the policies.

## 2.1 UNCERTAINTY IS EVERYWHERE

There are almost no policy issues that are not influenced by uncertain factors. Walker et al. (2013) argued that there are several reasons that considering uncertainty is essential in policy analysis. They explained these reasons by the hand of a couple different examples, of which some are briefly discussed here. Firstly, a research of Flyvbjerg et al. (2003) reviewed the cost and benefits analysis of multiple mega projects. It came to light that in nearly every case the costs were underestimated while the benefits were overestimated. Due to the usage of a single scenario or a small set of scenarios, the estimations were just one single value. If the future is uncertain, the future can not be portrayed by a single value, but should be a range of possible values to account for a wide set of different futures. Another example is found in the predictions of the growth of passengers and noise emissions by Schiphol airport. Researchers tried to predict the growth in passengers by following the growth of the Gross national product (GNP). For the GNP three scenarios were created, a low, medium and high scenario. The predicted growth of passengers was based on the highest growth of the GNP. However, in 1999 the GNP growth followed the line of the low scenario and the passenger growth was way higher than the growth of the GNP in a high scenario. What happened was that the passenger growth was also influenced by factors that were not known beforehand. As a result, the policies made for up until 2015 had to be changed. To conclude the main reasons to regard uncertain in policy analysis come down to this:

1. Uncertainty cannot be eliminated since we cannot see in the future.
2. Ignoring uncertainty can lead to poor policies and inefficient use of resources.
3. Ignoring uncertainty can limit our ability to take corrective action in the future

Beside the reasons of regarding uncertainty in policy analysis Walker et al. (2013) also describe 4 different locations of uncertainty in policy analysis:

1. Uncertainty in the external factors (X)
2. Uncertainty about the response of the system on external factors and policy changes (R)
3. Uncertainty about the outcomes of the model (O)
4. Uncertainty about the weights that are assigned to the outcomes(W)

Figure 2.1 shows in more detail how the uncertainties are located. External factor uncertainty (X) is based on the fact that we do not know how the future develops. Usually there is a prediction about the range of values the uncertainty can end

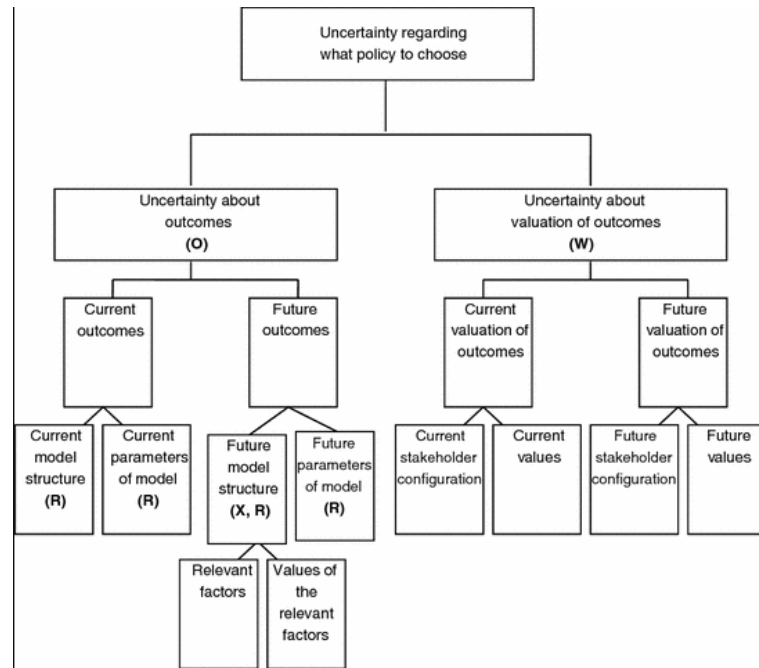


Figure 2.1: Locations of uncertainties in policy analysis (Walker et al., 2013)

in. Depending on the level of the uncertainty this range can be very small or very large. The relation uncertainty (R) is based on the fact that it is almost impossible to perfectly explain the (cor)relation between factors. This is most obvious in very complex systems (Agop et al., 2014). In complex systems there are a tremendous amount of variables that influence each other. Thereby there is the notion of emergence where the behaviour of the system cannot be explained by looking at single interactions. The example about the growth of Schiphol is a good example of a unpredicted relation. The outcome uncertainty (O) is the result of the two above mentioned uncertainties. Due to uncertain externalities and relations it is also impossible to give a single outcome value. The outcomes also have a range of different values based on the input that is given. The weight uncertainty (W) is based on the value that stakeholders give to outcomes. A stakeholder now can have a different view on life in ten years. For example, in the Dutch government the people in charge change every 4 years. The old minister of climate can have a more conservative view on global warming and will therefore weight less value to reducing CO<sub>2</sub> emissions.

## 2.2 WAYS TO DEAL WITH UNCERTAINTY

There are a lot of different methods to deal with uncertainty in policy analysis, ranging from a sensitivity analysis to Monte Carlo simulations. These methods are all viable in different systems and situations. The situation depends heavily on the uncertainties in the system that is analysed. By first analysing which uncertainties are present a useful method can be chosen. Methods like a sensitivity analysis are often only useful for dealing with level 1 or 2 uncertainties. Policy problems are usually influenced by many level 3, 4 and 5 uncertainties. Therefore, there are other tools needed to deal with these uncertainties.

Level 3 uncertainties are mostly dealt with by either reducing it to a level 2 or increasing it to a level 4 uncertainty. The first is possible by giving probabilities to values while the latter is possible by taking all the values as equal likely. For level 4 uncertainties it is best practice to find the most robust policy. This is a policy

that performs good in most of the plausible futures. This is also known as static robustness or scenario planning (Van der Heijden, 2011). Walker et al. (2013) discuss ways to deal with level 5 uncertainties. The first is resistance, where you try to plan for the worst possible future. The second is resilience, this means picking a policy that is able to quickly 'repair' the system after a disruption. The third and last one is adaptive robustness. The difference with static robustness is that you change the policy if the future develops different than is viable for the current policy. According to Kwakkel et al. (2012) the latter has proven to be the most fitting method. In later work Kwakkel and Haasnoot (2019) describe 5 other ways of dealing with level 5 uncertainty. These are Robust Decision Making (RDM), Dynamic adaptive planning (DAP), Dynamic adaptive policy pathways (DAPP), Info Gap decision theory (IR) and Engineering options analysis. The first is further elaborated on in next section.

### 2.2.1 Robust decision making

One of the most prevalent methods that take deep uncertainty into account is RDM. The main idea of RDM is to not use models and data as a predictive tool, but rather as a way to stress test policies on multiple possible future (Lempert et al., 2013). With the help of visualisations and statistics, analysts can then discover which uncertainties and policies are most influential in the futures that are of interest. This can, for example, be the future in which almost all policies seem to fail to meet a certain requirement. This way, policy makers are able to find a set of policies that meets multiple objectives. In other words, a robust policy.

#### *Foundations*

RDM is built upon a set of strong foundations (Lempert, 2019). The first foundation is decision analysis (DA). Empirical research has shown that people make better decisions when use decision aids. DA comprises the theory, methodology and practice of such aids. RDM applies this by evaluating alternative actions on alternative future states of the world and comparing trade-offs between these actions. The key difference between regular DA and RDM is that DA takes future states based on a probability function, while RDM regards future states as deep uncertain which does not include a probability function (Walley, 1991). In other words, DA focuses on optimality and RDM on robustness. This difference is also present in the approach. DA uses a predict-then-act approach (Lempert et al., 2004) and RDM uses an agree-on-decision approach (Helgeson, 2020). The former seeks the most plausible future to get to a set of alternatives. The latter tests a set of alternatives to get to a broader set of alternatives based on the possible future states of the world.

The second foundation of RDM is assumption-based planning (ABP). ABP aims to reduce the deleterious effects of over-confidence in existing systems and plans by improving understanding of how and why they may fail (Lempert, 2007). This is achieved by first identifying which assumptions are critical to make the plan work and then judge these assumptions on vulnerability. Based on the assumption that are both critical and vulnerable, three things are considered. Shaping actions, which try to make the assumptions less likely to fail, hedging actions, which can be taken if the assumptions start to fail, and signposts, trends and events that signal the start of a failing assumption.

The third foundation of RDM is scenario analysis. The idea of a scenario is that it is a way of looking at the future, without assigning a probability to it. In other words, it does not look at probability, but it looks at possibility. RDM uses scenario analyses to characterize and communicate deep uncertainty (Lempert, 2003). It applies quantitative scenario discovery algorithms to find a set of scenarios that are interesting. Then stakeholders can identify strategies based on these scenarios. Lastly, these strategies are tested over a wider set of scenarios to see which performs well overall. (Van der Heijden, 2011).



The fourth and last foundation of RDM is exploratory modeling (EM) (Bankes, 1993). EM is the concept that combines the three previous mentioned foundations. Exploratory models map a range of assumptions onto their consequences without giving a probability to these assumptions. This is effective if there is no consensus over the way these assumptions will develop or behave. With a proper experimental design EM can prove to be useful to inform policy makers about policy choices. It can help to generate a hypothesis, find interesting cases or tell something about the effect of different variables on the system. RDM makes use of all these benefits, but primarily uses the EM framework to find robust policies. Another advantage of EM is that it allows for global sensitivity analysis, since it does not use a base case as anchor point.

### Steps of a RDM

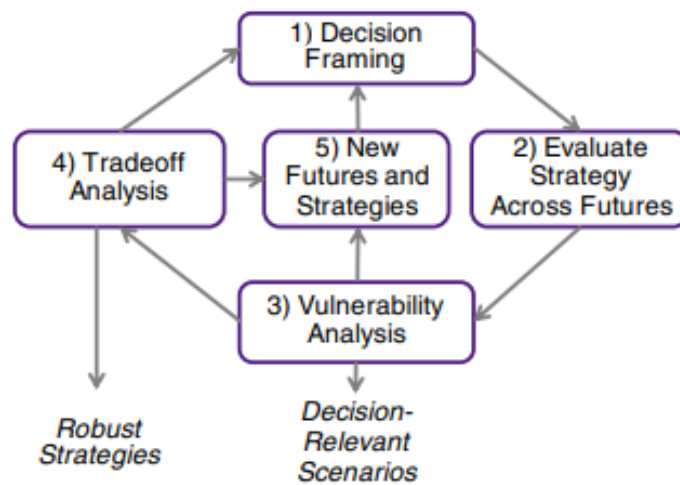


Figure 2.2: 5 steps of a RDM (Lempert et al., 2013)

The workflow of RDM, as can be seen in 2.2, is an iterative process in which every step can be reiterated based on the results in other steps. Below the steps are described in greater detail:

1. The model is defined based on a XLRM diagram (Lempert, 2003). These are the exogenous factors (X), policy levers (L), Relations (R) and metrics (M). The exogenous factors are the uncertainties in the model and have a range of values. The lever can either be some fixed policies or changing policies. In the case that they are changing, they are also given with a range of possible values. The relations describe how the system works and how the levers and exogenous factors are influencing the system. Lastly, the metrics are the outcomes of interest. These can be a time series in the case of a model that runs over time or one value in the case of a model that calculates for one point in time.
2. The second step starts with the selection of a set of interesting policies. This set can be chosen via a discussion with stakeholders, the public debate or by exploring the model itself. The latter can be achieved via a optimisation over the policy levers that are in the model. This is possible if there are a broad set of possible policies. When the policies are chosen, the policies are run over a wide set of different scenarios. These scenarios try to represent the uncertain future as good as possible. Every scenario is defined by a unique set of uncertainty values. In order to do this, techniques like Latin hypercube sampling (LHS) are applied. This method picks scenarios in such a manner that every corner of the uncertainty space is represented. The uncertainty space

is defined by all the combination between possible values of the uncertainties defined.

3. Visualizations and data analysis are used on the data retrieved from broad set of model runs. A prominent way to analyse the data in by doing a global sensitivity analysis. This shows the effect of all the uncertainties on the outcomes of the model. Another often applied method to explore and characterize vulnerabilities in the model is scenario discovery. This method searches for scenarios where the proposed solutions meet or miss their goal. The cluster of scenarios that are vulnerable can be used for further analysis. An adapted version of the Patient Rule Induction Method (PRIM) is incorporated in the EMA\_workbench as a scenario discovery algorithm (Hamarat et al., 2013).
4. In this step the analyst can make trade-offs between competing strategies. Since there are often competing objectives these trade-offs can help in balancing between these objectives. An example is cost vs reliability. You would like to maximize reliability while minimizing the costs. Problem is that lower costs usually mean lower reliability. The trade-off analysis can also be performed by looking at the score on robustness metrics between competing objectives.
5. After the vulnerability analysis or the trade-off analysis, new futures and strategies can be developed. From the vulnerability analysis it can become clear that a certain future causes the whole system to fail and the current policies cannot fix that. Than the modeller can go back to the decision framing to add other policies that do fix the problem in that specific future. Same goes for the trade-off analysis. If there are no policies in the analysis that perform above a certain threshold value, new policies have to be made.

#### *Latin hypercube sampling - RDM step 2*

LHS is a small-sample Monte Carlo approach. It preserves marginal probability distributions for every variable, while matching target correlations between variables (Huntington and Lyrantzis, 1998). LHS constructs a highly dependent joint probability density function for the random variables in the problem, which allows good accuracy in the response parameters using only a small number of samples. LHS consists out of two steps. In the first step a set of samples is chosen that represent the probability function of a input variable. In the second step the variables are ordered so they match the correlation between the variables. Due to the fact that correlations are enforced by the modeller the marginal probability distributions of the variables remain intact.

Figure 2.3 shows how LHS looks like in a two dimensional uncertainty space with 5 samples. Both variables are divided in 5 equal parts, which means that both variables are uniformly distributed. The 5 samples are chosen in such a manner that no row or column has more than 1 sample. This ensures that the uncertainty space is represented in a broad manner. LHS is also the default way of sampling in the EMA\_workbench, which is the main python library used in this research.

#### *Global sensitivity analysis - RDM step 3*

Global sensitivity analysis (GSA) is a important step in the assessment of uncertainty in mathematical models (Christopher Frey and Patil, 2002). GSA considers the influence of the full domain of uncertainties on the output behaviour. Therefore, the complete distribution of every input must be evaluated and in addition the importance of the input must be evaluated over the complete domain of the other parameters Liu and Homma (2009). GSA differs from normal sensitivity analysis due to the fact that GSA looks at the effect of the change of one single input compared to a base case. GSA is very helpful in complex cases like this due to the fact

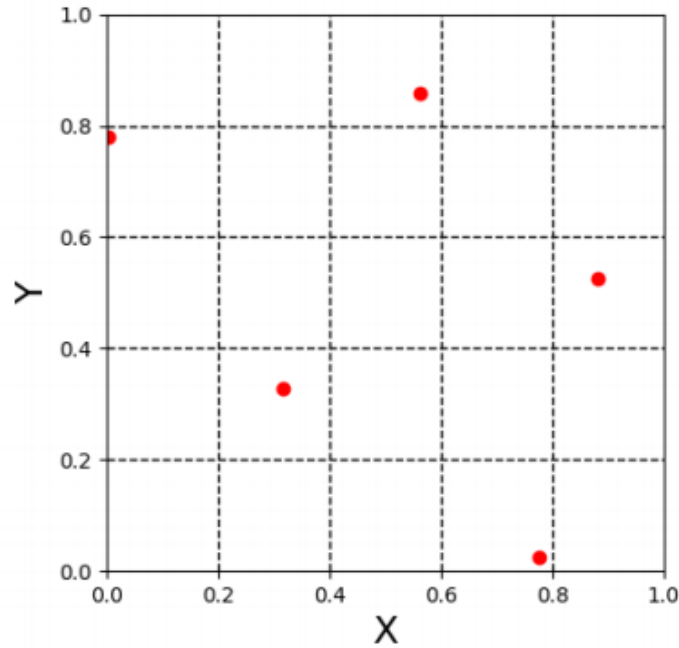


Figure 2.3: Example of a 2D LHS (van der Scheer, 2021)

that there are many different inputs with interactions that have to be considered. A downside to GSA is that it can become expensive rather quickly.

Saltelli et al. (2008) describes four different research objectives in regard to GSA. The first is factor prioritization, which finds inputs that have the most influence on the output uncertainty. The second is factor fixing, which searches for inputs that are negligible in their effect on output uncertainty and can therefore be taken as constant. The third is variance cutting, which looks at on which level of input values, the output values can be reduced below a threshold. The fourth and last research objective is factor mapping, which identifies spaces in the uncertainty space that produce certain outputs. Factor prioritization and factor fixing are the two research objectives that are of interested for the GSA in this research. This shows which inputs are important and are worth taking a closer look at and can reduce the size of the analysis by removing certain uncertainties in future runs. Factor mapping is part of scenario discovery and is therefore discussed in the next section.

There are many ways to perform a GSA. Jaxa-Rozen and Kwakkel (2018) discuss the effectiveness on decision trees in performing GSA. It tries to find the splitting criteria between a set of input combinations, and regions of the output space. Individual trees usually show a lot of variance, which means that a small change in the input data can transform the structure of the tree. Therefore, techniques are developed that aggregate a bigger set of different trees. One of the most prominent is extra-trees (Geurts et al., 2006). This algorithm produces more accurate results and is also computational efficient.

The EMA\_workbench has a certain method to implement extra-tree algorithms to gain insight into the effect of uncertainties on outcomes. This method is feature scoring (Kwakkel, 2021). Feature scoring is often used in machine learning to find the most relevant factors. It is very efficient to use for factor prioritization in GSA. An advantages of this technique is that it can also handle categorical factors like policies and therefore it does not put any constraints on experimental design.

### *Scenario discovery - RDM step 3*

Scenarios are tools to communicate uncertainty in decision making. However, choosing scenarios to run the model for can be very challenging in complex models with a wide range of uncertain dimension. [Lempert et al. \(2008\)](#) compare different methods to identify scenarios of interest, which is also referred to as scenario discovery. They compare the classification and regression tree (CART) method and the patience rule induction method (PRIM) method. In this section only PRIM is explained further. The goal of scenario discovery is to find a set of input parameters that are strongly predictive to a certain set of outcomes. The idea is that one or multiple criteria are given which the outcomes must suffice. The outcome of the scenario discovery is a region in the scenario space for which this is true.

The first method compared is based on PRIM, which is a method brought forward by [Friedman and Fisher \(1999\)](#). [Lempert et al. \(2008\)](#) use PRIM to find boxes within the uncertainty space which contain a high level of policy relevant scenarios. These boxes should give the user a simplified image of the relevant inputs within the uncertainty space. There are three metrics that are used to measure the quality of the boxes. The first is coverage, which looks at the percentage of true case in the box compared to the total amount of true cases. The second is density, which looks at the percentage of true cases in the box compared to the total amount of cases in the box. The last one is interpretability, which looks at how easy the visualizations can be understood. This is usually based on the amount of boxes that are in a box set which explain the data.

PRIM is an iterative process with two main steps, which can be found in appendix C figure C.1. It begins with the entire data set and no restricted dimensions. Then it creates a series of boxes that are increasingly smaller and denser. PRIM finds these smaller and denser boxes by removing a small slice which is picked based on the highest increase in density. The series of boxes that come out of this so called 'peeling trajectory' are then visually presented to the modeller. The modeller can pick a certain box that has the best balance between coverage and density according to the modeller. After the box is chosen the boxes can be expanded via pasting. This is a process where unnecessarily restricted dimensions are allowed to expand. The next step in the process is called covering. The data within the chosen box is deleted from the data set and the peeling/pasting process is repeated. These steps can all be reiterated until the modeller decides the algorithm cannot adequately find useful boxes.

### *Robustness metrics - RDM step 4*

The trade-off between the competing strategies is often based on a robustness score for the outcomes, which depends on a robustness criteria ([Hall et al., 2012](#)). This criteria determines how robustness is viewed in a certain research. A few examples of robustness criteria are the signal-to-noise ratio which takes the mean of a dataset divided by its standard deviation or minimal regret which takes for every scenario the difference between the best performing scenario and the scenario reviewed. Both methods usually generate roughly the same outputs. If for example the best scenario is an outlier compared to the other scenarios, the signal-to-noise ratio will, on average, have a better value than the minimal regret has. This is because this one scenario makes the regret on all other scenarios automatically way higher. The choice for a robustness metric depends on the preference of the researcher. Another option is to apply multiple metrics at first and choose one as leading afterwards.

The robustness metrics used in this research are a set of the proposed metrics by [McPhail et al. \(2018\)](#), which are also found in figure 2.4. They scale the metrics from low risk aversion to high risk aversion. A high risk averse metric weights bad outcomes highly while low risk averse metrics tend to mainly look at the good performing outcomes. Robustness metrics have two more ways in which they differ [McPhail et al. \(2018\)](#). The first is if a metric looks at absolute or relative performance.

The second is if a metric tries to satisfy a certain condition or if it looks at actual system performance. In this research the set that is used encapsulates as many dimension as possible.

The used set is listed below.

- **Minimax regret:** This is one of the most risk averse metrics. It looks at the relative values between policy outcomes of all scenarios. The idea is that for every scenario every policy is compared to the policy that performs best, this is called the regret. After that, the minimax regret of a policy is the highest regret the policy has over all the scenarios. The benefit of this metric is that you can reduce the change that you pick a policy that is the worst in a certain scenario. However, this metric is very susceptible for outliers. One outlier for a policy can make the minimax regret very high, while the average regret might be a lot lower.
- **90th percentile minimax regret:** This metric fixes the problem that minimax regret might have. Instead of take the maximum regret, this metric picks the 90 percentile regret. If in 10% of the scenarios the policy performs poorly it is not an outlier anymore. A negative side is that it is still possible that the outcome falls within the worst 10%. This is something that has to be considered by the decision maker.
- **Maximax:** This is the least risk averse metric thinkable. Maximax purely looks at the scenario in which the policy performs best. This is very easy to calculate, which makes it easy to apply. However, the big problem is that over a set of 100000 scenarios this one good outcome says very little about the overall performance. This metric is not used often in decision making under deep uncertainty. A danger is that the modeller uses that one good performing scenario as the outcome to show to the decision maker. This way, on average bad performing policies are implemented due to a single good performance.
- **Mean variance:** This is the same metric as the signal-to-noise ratio of the previous paragraph. It is a metric based on absolute performance and the level of risk is on a medium level. As mentioned, this metric takes the mean divided by the standard deviation. The benefit of this is that an on average poor performing policy can still proof good due to their consistency.
- **Starr's domain criterion:** This is a satisficing metric. The idea is to find the volume of scenarios in which a certain criteria is met. According to this metric, the policy with the highest volume of correct scenarios is the most robust. On the one hand, this metric gives you the policy with the highest change of meeting your goal. On the other hand, this metric does not say anything about the negative consequences a policy can have if the goal is not met. Thereby, this metric is vulnerable to uncertainty from the side of the decision maker. The decision maker can slightly lower or higher the criteria to get out the policy or their liking.

## 2.3 CONCLUSION METHODOLOGY

One of the prominent methods to deal with deep uncertainty is RDM. With this method the most robust policies for the uncertain future can be found. There are 5 steps in the RDM cycles. The first step, decision framing, is performed in the implementation part of this thesis. The second part, evaluating strategies over futures, is setup and performed in the first part of the results. The experiments are drawn via LHS. The third part, the vulnerability analysis, is also performed in the results section. A global sensitivity analysis is conducted with feature scoring. This

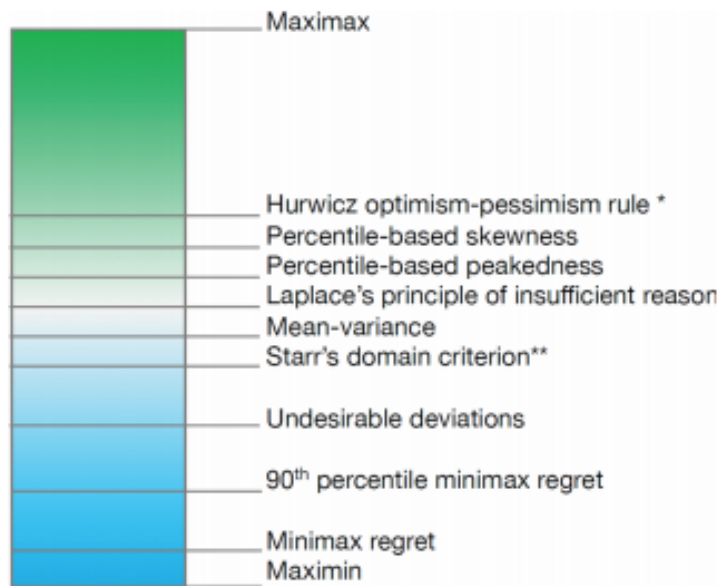


Figure 2.4: Robustness metrics as defined in McPhail et al. (2018). The bar goes from green to blue, where green is a low level of risk aversion and blue is a high level of risk aversion

shows the relative effect of all the uncertainties on the output variables. After that, scenario discovery shows some regions in the data which produce results of interest. The PRIM method is applied in order to do this. The fourth step of the RDM cycle is also the last part of the results. Here, five different robustness metrics are compared. These metrics all reward policies in a slightly different manner. The last part of the RDM cycles will only be performed if new futures and policies are deemed interesting.



# 3

## MODEL IMPLEMENTATION

In this chapter first step of the RDM cycles, which is the decision framing, is performed. This means that the XLRM diagram is defined. These are the externalities, levers, relations and metrics. Thereby, the Groningen model has to be adjusted to have a lower run time and the connection between python and Visum has to be made. First the latter part is discussed and after that the XLRM diagram is defined.

### 3.1 ADAPTED MODEL

As mentioned, the Visum model has a run time that is too high to apply to the EMA\_workbench and must therefore be reduced. Reducing the run time is an iterative process. With every step, more is learned about the model. The most important steps and changes are described in greater detail.

#### 3.1.1 Considerations to make when adapting the model

There are some important considerations to make when reducing the run time of the model. Firstly, the model must be reduced in a way that it still suffices the purpose of this research. The purpose of this research, in regard to the model, is to run a sufficient amount of different scenarios with the model to get the distance travelled by car within the municipality in every scenario. This means that the amount of scenarios has to be determined and that the distance travelled by car within the municipality must be produced by the model. This introduces the second consideration, the outcomes have to be valid. Changes in the model can cause big changes in the results of the model. This can be checked in several ways. The most time-efficient way is to compare the produced data from the adapted model to the data produced by the original model. By doing this the expectation of experts is compared to the model of this research. Landry et al. (1983) refer to this as "convergent validation". The third and last consideration is that the run time should be manageable for the amount of runs. Optimally the run time would be not more than 1 second. However, the model structure does not allow such a big reduction. Regarding the fact that the planned time for experimentation is 3 weeks, the total run time should not surpass the duration of a week. Later in chapter 4 a more accurate approximation is made.

#### 3.1.2 Changes in the model

Initially there are two main parts of the model that can reduce run time. The first is in the model network. This could, for example, be the level of detail of the streets or by aggregating the zones found in the model. The second is in the procedure sequence. Here the amount of procedures can be reduced or some of the heavier procedures can be removed. Another option in the procedure sequence is reducing the amount of iterations done in the route assignment and the mode choice part. Table 3.1 describes these steps.

The three steps that have reduced the most time are the removal of the route assignment part, the reduction in the MD model and the reduction in iterations



Table 3.1: Steps taken to adapt model

	Model part	Change	Reason	Gone through?	Time reduction
1	Network	Reduce Zones	Amount matrix calculations is zones <sup>2</sup> . Reducing zones reduces computation time.	No. Reducing the model structure would have needed too much work, because things like demographic data are also entangled with this.	-
2	Procedure sequence	Reduce iterations assignments and mode choice model.	An assignment or mode choice model does the amount of iterations it needs to find an equilibrium or until the max amount of iterations is reached. Reducing the amount of max iterations	No. Model time was reduced just a little while the results where less accurate. The trade-off between time and quality was not worth considering.	-
3	Procedure sequence	Remove route assignment	Takes the most time and is not needed to determine the modal split.	Yes.	5 hours
4	Procedure sequence	Reduce MD model steps	The first step, of 15 steps, of the MD model calculates the mode choice for 80% of the tours.	Yes.	2 hours 5 minutes
5	Procedure sequence	Creations of tables	Originally the MD model creates 4 end tables, from this tables only 1 is useful for the purpose of this research.	Yes.	3 minutes
6	Procedure sequence	Reduce model iterations	The total MD model usually takes a weighted average of 3 iterations. Since the goal of this research is to run at least a broad set of different scenarios, this weighted average is not needed.	Yes.	No. Amplifies all time reductions by a factor 3

of the model. To give a short overview the MD model first took around 2 hours and 15 minutes and the route assignment part around 5 hours. If this is done for three iterations this equals 21 hours and 45 minutes. The reduced model only takes around 12 minutes, where 9 minutes are taken by the first part of the MD model. Other procedures like changing attributes and changing matrices account for the 3 minutes that are left.

Adapting the model is a very fragile process. Small changes can be devastating for the quality of the model. Therefore, changes have to be made with caution. Changes in the model network seem to have a disastrous effect. A small change in the network results in outputs being completely unviable. This is a con of having a model with such a high level of detail. However, changes in the procedure sequence do not have this extreme impact. These changes have a big impact on the quantity of data that the model produces. Removing the route assignment part does leave out data about how tours are exactly taken, but it does not effect the mode choice part at all. Therefore, all the changes in the adapted model have been made in the procedure sequence.

The changes in the model all have some effect on the performance of the model. The first change is the removal of the complete route assignment part. This has no effects on the distance travelled by car within the municipality, which is the interesting output of this research. The difference with the old model is that there is no information about how exactly traffic is divided within the municipality. However, this is mostly useful for insight into congestion on roads. The second change is the reduction in MD model steps. Initially there were 15 MD steps in the model. The first calculates all direct tours and the other 14 all the indirect tour. Direct tours are for example tours from home to work and back home. Indirect tours are for example from home to work to the supermarket and back home. So there is at least one extra stop within the tour of an indirect tour. In the changed model all indirect tours are removed from the procedure sequence. The indirect tours account for approximately 20% of all tours. This means that the results of this research are based on only 80% of all tours. These results are validated in the next section. The third alteration is the amount of times that the route assignment and mode choice is performed. The model originally runs this three times which is changed to one. The model ran three times to account for random deviations in the results. However, in the light of deep uncertainty running three times does not cover the randomness. Therefore, changing this to one iterations does not change the model outputs much. The last change in the model is removal of some tables that can be made. Since the removed results are not used in the analysis or the model itself, it has no effect on the outcomes of interest.

### 3.1.3 Validation of the adapted model

In order to validate the model a convergent validation method is applied. This method compares the results to the expectation of experts. In this case, the results from the original model is based on a expert opinion. Therefore, the results of the adapted model are compared to that of the original model. The outcomes that are looked at are: the total distance travelled, the share of the car in the total distance travelled and the total amount of distance travelled by car. These outcomes are chosen due to the fact that the municipality of Groningen wants to know how much distance is travelled by car. The first two determine the third, which is the most important outcome of this research. The original model is run 4 times and the adapted model 10 times. There are multiple runs to account for some deviations in the results. The second model has more runs because the run time allowed for a higher amount of runs. For all outcomes the average of the runs for each model is taken.

Figure 3.1 shows the comparison of the results of the original and the adapted model. As mentioned, the expected change in distance travelled was approximately 20%. This expectation seems to be correct. This change in total distance travelled does not have to influence the outcome of this research. This is the case when the distribution over transport modes stays the same. Therefore, it is important that the car share does not change. This is possible when direct tours are for example taken by car more often. The difference between the two models in car share is not more than 1%. Since this is a very small change compared to the original model, this difference is negligible. It is to be expected that the same interaction found in the original model are found in the adapted model. Due to the fact that the total distance travelled changes by 20% and the car share is constant, the total distance travelled by car also changes by 20%. This shows that the model correctly calculates the total distance travelled by car.

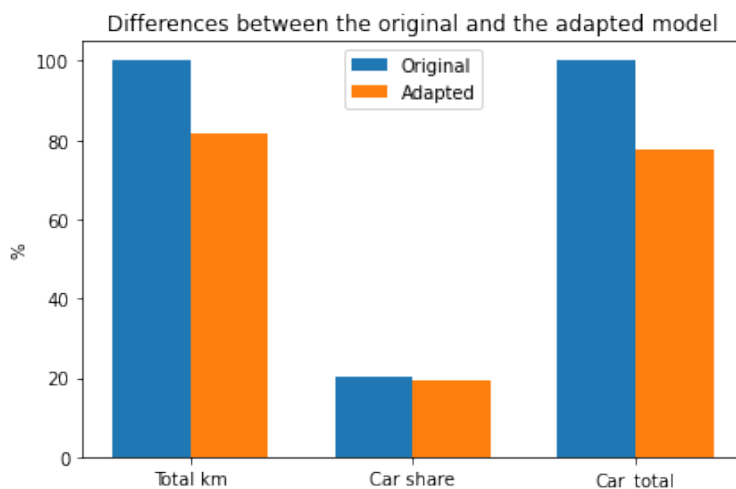


Figure 3.1: Comparison of the original and adapted model based on three different outcomes

## 3.2 CONNECTION BETWEEN VISUM AND THE EMA\_WORKBENCH

One of the prevalent tools of doing a RDM study is the EMA\_workbench. This is a python based library which is specially developed for exploratory modelling and analysis (EMA) research. The workbench hands tools to run a wide set of scenarios on a model. The user has to define a model with its levers, uncertainties and outcomes. Than the workbench draws scenarios from a range possible values for all uncertainties and levers. For example, gasoline price is predicted to be between

€0,80 per liter and €1,20 per liter by 2030. The workbench takes a different value between the upper and lower bound for every scenario. It takes a certain method to draw these values, this can be LHS, Monte Carlo sampling or a full/partial factorial design. Depending on the amount of runs that can be done, one method performs better than the other. As mentioned in chapter 2, LHS is used in this research. The model is then run with the different scenarios as input. The workbench offers simple to use tools to present the retrieved results. PRIM and feature scoring are also both provided by the EMA\_workbench.

In order to apply to functions and tools of the EMA\_workbench to the Visum model a connection between the two is made. Visum already has a built-in python console and is therefore also very easy to connect to python. Visum does this connection via a COM interface. A COM interface allows for the decoupling of the method and the implementation. Figure 3.2 shows how this conceptually works. It also shows which RDM steps are represented by the different parts. From the top left corner, it all starts with defining the inputs. These are the uncertainties, levers, the outcome and most importantly the model it self. With the possible input values the EMA\_workbench generates a set of scenarios based on LHS. The Python script now iterates through the list of scenarios. After every iteration the information of the scenario is sent to the Visum model via a COM interface. The Visum model applies all the newly received input values and run the procedure sequence. The Visum model then communicates, via a COM interface, with the python script to start a new scenario iteration. After the model has ran all scenarios the data is ready to be analysed. In this case a vulnerability analysis and a trade-off analysis can be performed.

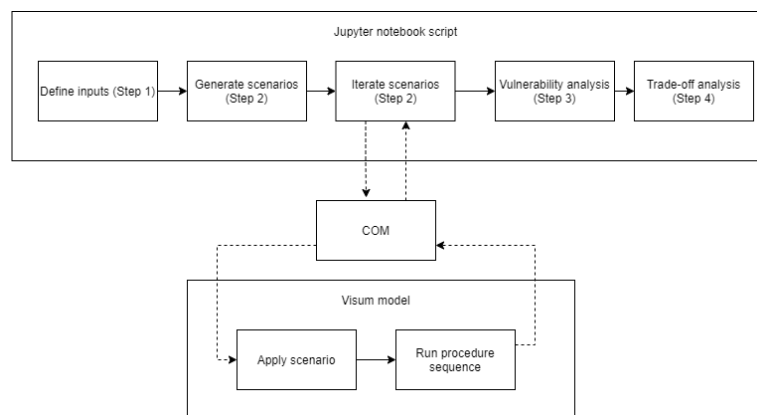


Figure 3.2: Conceptual connection between the EMA\_workbench and Visum

The code which implements the above concept can be found in <https://github.com/ilmovanbaarle/Thesis-Ilmo>. The code in `visum_connector.py` is the code that perform both step 1 and step 2 of the RDM cycle, as shown in figure 3.2. This means that it also connects the Python model to Visum. Firstly, the model is defined. This model is a child of the parent class "VisumModel". The parent class "VisumModel" has some interesting features. However, to understand these features first the inputs of the model have to be defined. These are the uncertainties, outcomes and policies. The first two are given in a range while the policies are all just one single policy. The policies have a "model\_file" as attribute. This "model\_file" defines which model is run. Each model has a different policy active. Then the results are defined by the performs of a given set scenario runs. These scenario are picked via LHS from the ranges given for the inputs. Every scenario is put into the "VisumModel" child. In the "VisumModel" class the model is initialized. Via a COM command Python and Visum are connected and with the help of "model\_file" the correct model is opened. Now that the model is opened, the uncertainties can be changed. The "Visum.Net.SetAttValue(k, v)" changes all uncertainties in the Visum model. With

"Visum.Procedures.Execute()" the model performs a run. Finally the results can be retrieved from the Visum model. Beforehand, Visum is given the task to create an excel file with information about the mode choice in every region of Groningen. The "output\_changer" function then turns this data into the outputs of interest. To reduce the amount of opened files on a single computer, the "VisumModel" ends with closing the Visum model. Closing the model only happens if a different policy is run, if the same policy is applied in the new scenario there is no need to change the "mode\_file" since this only changes the policy. This also reduces the amount of times that the model has to open and close a "model\_file". The Python file repeats this sequel with a different scenario, until all scenario are run. With the retrieved results the other 2 steps can be performed within Python.

Since there is a wide set of scenarios that are run, running multiple Visum models simultaneously would reduce total run time. There are 2 options handed by Visum to perform this. The first is to distribute the model runs over different computers. This way the run power of a single computer is not exhausted when multiple runs are performed on the same computer. In theory one could couple 10 different computers through Visum. The second option is to run multiple runs on a single computer. Visum allows for 5 parallel runs on a single computer. This is due to the license that is given out and the possible memory capacity problems. In this research the second option is only used. This means that instead of 1, 5 models are run in parallel. This does increase the run time of a single run slightly. Due to problems in memory usage, 95% - 100% is constantly used, the computer has a little more struggles when computing. The run time increased from around 12 minutes to around 15 minutes. The first option is not used due to the more challenging coupling process and the already sufficing second option. However, a combination of both options would give some great scaling opportunities. If only 5 computers would be connected, 25 models could be run simultaneously. This is something to consider when more resources are available.

### 3.3 STEP 1: DECISION FRAMING

The inputs of the EMA\_workbench are based on a XLRM-diagram. An XLRM-diagram contains externalities (X), levers (L), relations (R) and metrics (M). The externalities are also referred to as uncertainties, the levers to policies and the metrics to outcomes. These things mean the same, but for the sake of comparison with the Groningen model the names are changed.

#### 3.3.1 Externalities

The Groningen model is influenced by a set of uncertainties. In the original scenarios that are run, two values for uncertainties are taken. These values are determined on the high and the low WLO scenarios ([Planbureau and voor de Leefomgeving, 2015](#)). Table 3.2 gives the names of the uncertainties, the values that are taken from the WLO scenarios and the adjusted values. The idea is to take the low scenario as lower or upper bound and the high scenario as the lower bound or upper bound. It has to be noticed that these values are from 2015. Anno 2020-2021 we are dealing with COVID-19 which has shown some new insights. Especially the amount of people working from home have grown a lot during the pandemic and expectations are that some of this growth is detained after the end of the pandemic ([van Eck et al., 2020](#)). Therefore, the range of reduction in tours due to people working from home is slightly increased. Due to the fact that the developments are very recent, a scientifically substantiated value is not yet available. The adjusted values are based on a personal expectation which is possible due to the deep uncertain origin of this value. The public transport costs per km are also adjusted. According to the WLO

Table 3.2: Uncertainty ranges

Uncertainty	Acronym	Low scenario	High scenario	Adjusted range
Normal E-bike use	EBIKE.BASIS	0,220	0,280	-
E-bike used for education	EBIKE.OW	0,090	0,110	-
Car costs per km	KMKOSTENINDEX	0,953	0,700	-
Public transport costs per km	OVKOSTENINDEX	1,100	1,100	1,000-1,200
Reduction in tours due to people working from home	THUISWERKREDUCTIE	1,000	0,950	0,850 - 1,000

scenarios this value will not change. However, technical developments or national policies can still have a big impact on the costs per km for public transport. The costs per km for public transport can be regarded as a deep uncertainty and therefore the new upper and lower bound are not scientifically grounded. In order to clearly see the effect around 10% above and 10% below the WLO values is taken. All other uncertainties are not changed due to the fact that there is no reason to assume that these values end up differently.

### 3.3.2 Levers

The levers or policies are the actions the municipality can take to influence the system. Sweco received a table of policies from the municipality. This table is presented in figure A.1. They present 6 different policies with a high scenario, a low scenario and a high scenario without Oosterhamrikstracé (OHT). The high scenario presents a higher intensity of a certain policy.

The policy that are chosen are listed and described in detail:

- No policy: To see the effect of a policy clearly there is also a policy that implements nothing. There are situations where the base case performs even better than the non-base case. It is good practice to always include a base case scenario.
- Bike policy: This policy is about increasing the amount of so called "snelfietsbanen", which in translation means lanes where people can bike fast. This policy focuses mainly on the expected increase of E-bikes. The higher speed of E-bikes cannot fully be used if there are many other types of traffic or if there are many side lanes. The "snelfietsbanen" secure a more free road for all cyclists. There are some roads in the Groningen model that are certainly changed into "snelfietsbanen". Another group of roads is only changed into a "snelfietsbaan" when this policy is active. The main effect of "snelfietsbanen" are higher speeds for cyclists.
- 30 km/h policy: The idea of this policy is to reduce speeds on roads within the city of Groningen from 50km/h to 30km/h. On the one hand this policy can reduce the number of and severity of accidents. If for example two cars hit each other with 50km/h, the damage is worse than if two cars hit each other with 30km/h. Next to the accident argument, this policy focuses on reducing emissions of cars due to lower speeds and it focuses on reducing the attractiveness of taking the car. The reason that the attractiveness reduces is that, if the speed reduces the time to get somewhere in the city also takes longer.
- Parking policy: There is a small area in Groningen where you have to pay €4/h to park your car and a bigger area around that where you have to pay €2,40/h. In the parking policy the municipality wants to increase the size of the second area. So places where there are no parking costs as of this moment will get a parking cost of €2,40/h. This makes it less attractive to own a car in the city and also to travel to the city by car. In the high scenario there is a bigger area that gets parking costs than in the low scenario. For this research the high scenario is applied.

Table 3.3: Ranges of changing policies

Lever	Variable name	Upper bound	Lower bound
Speed of bikes	BIKE_SPEED	25	15
Speed of cars	CAR_SPEED	50	30
Parking costs	PT_KP	500	0

### 3.3.3 Changing levers

The EMA\_workbench gives the possibility to also change the level of policies. From a policy maker perspective it is very interesting to see what the effects of different levels of policies are. There are multiple benefits to this. The first benefit is that the policy maker gets a more clear view on the influence of a policy on the outcomes. There can be a certain threshold value for which the policy gets a different effect. For example, the effect of the price of parking on car use. It can be that between 0 and 1 euro the effect is linear and after that the effect is exponential. A second benefit is that the effects of different policies together can be researched. This hands the policy maker a broader set of combinations of different policies that perform well. This also shows the interaction effects between policies. A downside is that the levers are part of the sampling. This means that more runs are needed to present the uncertainty space in the same manner as with a single value policy. In other words, it is more computational expensive.

The three policies used for the changing policies are the same as the policies used for the non-changing policies. The idea is that the three policies are always active, but the levels of the policies change for each scenario. The variables and values are found in 3.3. The bike policy is put in a variable called BIKE\_SPEED. Normally, when this policy is applied there are fast bike lanes on which bike can have a higher speed. However, the speed a cyclist can get to depends mostly on the degree of freedom on the lane. If there are very little places to get on to the lane or it is completely separated from the normal road, the average bike speed will be higher. Therefore, this variable can change between 15 km/h, which is equal to original speed without a fast bike lane, and 25 km/h, which is equal to the speed with fast bike lanes plus the difference between the speed with and without fast bike lanes. The second variable represents the 30km policy. If this policy is enforced by speed bumps or regular speed checks people shall drive 30km/h. However, if solely a sign with 30km/h is placed people are more likely to drive a bit faster. Therefore, this policy has as a max value of 50km/h to represent people not following the rules at all. The third variable is parking costs. The policy changes the parking costs for the whole parking area to a value between 0 and 5 euros. The upper bound is chosen at 5, due to the fact that it is not much higher than the 4,20 euros they charge now and it is a easy value to work with. The lower bound is equal to removing all the parking zones. This includes the 4 euro/hour zone and the 2,40 euro/hour zone.

Changing policies does change the way policies are implemented in the Python model. The policies are now regarded as levers and can change in value. This means that they are also taking into the sampling of the scenarios. The benefit of this is that the uncertainties are more widely represented. The reason for this is that they can take on as many value as the amount of scenario that are run, while previously the amount of different values was equal to the amount of scenarios divided by the amount of policies. Operationalization wise, the fact that the all policies are always active means that there is no need to close model files between runs. This reduces total run time slightly.

### 3.3.4 Outcomes

One of the most important considerations in the model is the set of outcomes you take into account. In this case the outcomes are chosen based on the goals of the

Table 3.4: Metrics regarded in the model

Outcome	Variable name	Unit
Total amount of movements	total_verpl	Movements
Percentage movements by car	carshare_verpl	%
Percentage movements by bike	bikeshare_verpl	%
Percentage movements by public transport	OVshare_verpl	%
Total distance travelled	total_km	km
Percentage distance travelled by car	carshare_km	%
Percentage distance travelled by bike	bikeshare_km	%
Percentage distance travelled by public transport	OVshare_km	%
Total distance travelled by car	total_car	km

municipality of Groningen, the available data and the policies used. The MD model produces two tables: a table that contains the amount of tours per mode per travel from every region to every other region purpose and a table the contains the amount of kilometers traveled per mode per travel purpose from every region to every other region. The metrics chosen are based on these two tables and are found in table 3.4

The last outcome from table 3.4 is the most important. Since the goal of the municipality of Groningen is to reduce the amount distance traveled by car in the municipality. However, the other outcomes can help in understanding how tours are divided between the different modes of transport. This is useful if for example the municipality also want to encourage people to take the bike, instead of the public transport. Thereby, the total distance travelled and the percentage of that distance travelled by car can help to understand how the total amount of distance travelled by car is built up. This can show on which aspect of the outcome the policy works. For example, more people working from home should lower the amount of distance travelled by car. However, this is not caused by less attractive car use, but by the fact that there are less travels in total. It is important that this relationship is understood correctly, because this helps the municipality in understanding the effects of their policies. Looking at the data set, the data set divides the movements based on region. These regions are divided as showed in table 1.1. For the outcomes only region 1, Municipality Groningen, is used for the results.

### 3.4 CONCLUSION MODEL IMPLEMENTATION

The first part of this chapter discussed the run time of the Visum model. The RDM method requires a big amount of scenarios. Therefore, the run time of the Visum model is reduced. The model ran for 22 hours and 45 minutes, which is now only 12 minutes. The biggest time reductions are gained due to the removal of the route assignment, a reduction in the amount of iterations and a big reduction in the amount of MD models. All changes are made while preserving the purpose of this research, getting the distance travelled by car within the municipality. The second part discussed the connection between Visum and Python. This connection is made via a COM interface. Thereby, Visum allows for 5 simultaneously running models,

which reduces the total run time. The third part showed all the uncertainties, levers and outcomes for this research. A total of 5 uncertainties, 4 levers and 10 outcomes are defined. The most important outcome being the total distance travelled by car within the municipality. From the lever side the decision is made to do some fixed and some changing policies. With this additional part it is possible to see how sensitive the outcomes are to the policies. This could convince the municipality of Groningen to adjust some proposed policies.





# 4 | RESULTS

The result chapter has four main sections, which describe the second, third and fourth step of the RDM cycle. The first part is the experimental design. Here it is determined how the experiments are drawn and how many experiments are drawn. The second part is the global sensitivity analysis of the model consisting of feature scoring. The third part includes the scenario discovery which is done with the PRIM algorithm. The last part dives deeper into the robustness of the policies, where multiple robustness metrics are compared.

## 4.1 STEP 2: EXPERIMENTAL DESIGN

The second step of the RDM cycles is evaluating strategies across a set of futures. Since there are two sets of policies, some that change in value and some that do not, two sets of experiments are run. Every scenario consists out of a different combination of values for uncertainties and levers. The value all fall within the range of that variable. As mentioned in chapter 2 the values for the variables in the scenarios are drawn via LHS sampling. The values that the uncertainties and levers can take on are found in chapter 2.

There are two important consideration to make when deciding on the amount of experiments run. The first is that the scenarios should fill the uncertainty space and lever space in a broad sense. There is no real consensus on when the spaces are filled enough. The second consideration is that the run time is in the time frame of the research. Running a model for multiple weeks would be to long. Taking these two consideration and the fact that the model runs for approximately 12 minutes, a decision has been made to run a set of 1000 scenarios for each of the policy sets. This makes the run time for both approximately 3 days, which means that both can be run in 6 days. For the non-changing policy set this means that every policy runs 250 different scenarios. These 250 scenarios do not change across different policies. In a 5-dimensional uncertainty space this might be too little. If it does not become clear what the relations between variables are from these 1000 scenarios, more scenarios can be run. For the changing policies this means that a 1000 scenarios are run in a 9-dimensional space. Again, this might leave some important corners in the uncertainty and lever space covered. However, if there is any doubt that the relations between variables are unclear or that there might be a combination of uncertainties that is worth to consider, more scenarios are run.

## 4.2 STEP 3: GLOBAL SENSITIVITY ANALYSIS

A global sensitivity analysis gives insight into the effect of uncertainties and levers on the output variables. Additionally, with the use of the PRIM algorithm regions of scenarios that are of interest can be found. First the effects of all the uncertainties and policies on all the output variables are discussed via feature scoring. Then the feature scores per policy are shown and discussed to see what the effects of policies are on the system. After that, the PRIM analysis uncovers some scenario regions of interest.

#### 4.2.1 Feature scoring

Feature scoring is a mix of techniques used in machine learning to find the most influential factors (chapter 2). The main idea of feature scoring is to look at effect of an input variable on an output variable. The value lies between 0 and 1 where 1 means that it completely defines an output variable and 0 means that it has no effect. The columns always add up to 1, so the values are relative to each other. This means that for example the effect of *THUISWERKREDUCTIE* on *totaal.verpl* can be very small (figure 4.1). However, due to the fact that the effect of the other variables is even lower it still accounts for 78% of the changes in this output variable. The most interesting values are discussed below.

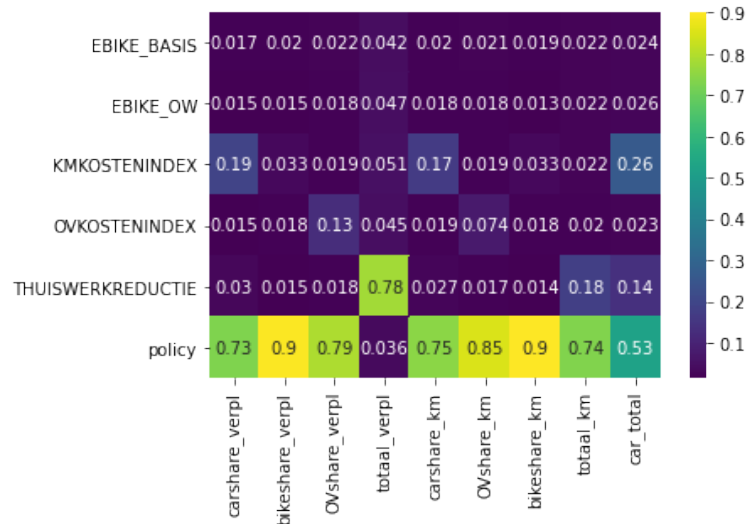


Figure 4.1: Feature scoring with policies

The first thing that is apparent from the feature scoring, in figure 4.1 is the high value of *THUISWERKREDUCTIE* on *totaal.verpl* and the way lower value on *totaal.km*. However, the policy score does have a way bigger impact on *totaal.km*. There are multiple explanations possible for this effect. One logical effect can be that if people work from home more often the amount of short movements decreases a lot, but the amount of long movements does not change a lot. The fact that the policies do effect the amount of distance travelled but do not really effect the amount of movements is difficult to pin down.

The second thing that is apparent is the high influence of policies on the outcome variables. Around 80% of the changes in movements and distance travelled for all the transport modes can be explained by the policies. This means that the policies have a big influence of the system. It is not yet clear from this feature scoring which policies influence the outputs and if they influence them in the right direction. One of the next paragraphs researches this.

Due to the high effect of the policies, it is also interesting to see what the effect is of the uncertainties when the policies are not taken into account. This is done for every policy. This does mean that the sample size goes from a 1000 to 250 for these 4 analyses. Figure 4.2 shows the results for the base case. The first take is that for the main kpi, which is *car\_total*, the uncertainties *KMKOSTENINDEX* and *THUISWERKREDUCTIE* are the main influence. The former determines around 60% of the output and the latter around 30%. The other outputs mostly correlate with the uncertainties in an expected manner. The costs for public transport has the highest influence on the movements and distance travelled with public transport and if more people own an e-bike there are more movements and distance travelled per bike.

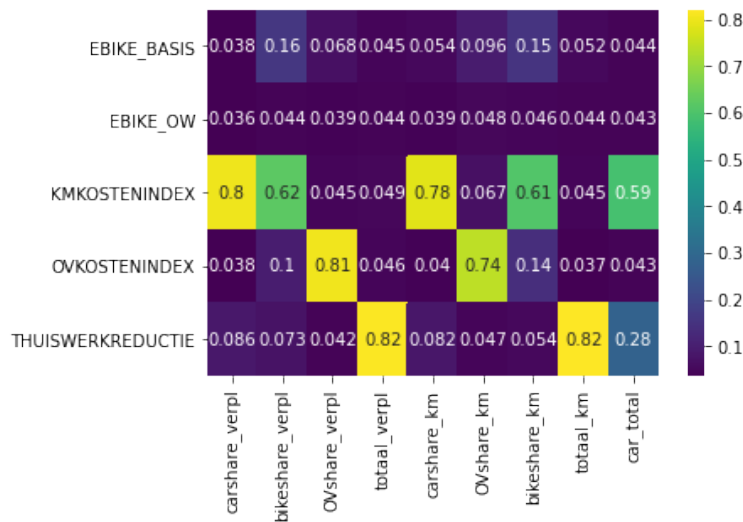


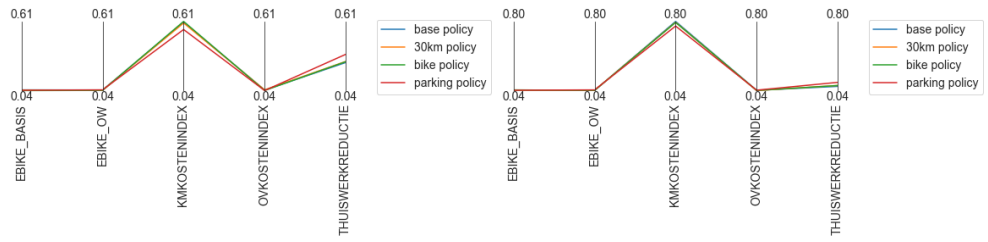
Figure 4.2: Feature scoring in the base case

Three outcomes are interesting from this feature scoring of the base case. Firstly, the uncertainty that has a high impact on the amount of car movements, also has a high impact on the amount of bike movements. Thereby, the effect on public transport movement is almost none existent. This could mean that the shift from car to bike (and the other way around) happens more often than the switch from car to public transport. Secondly, the uncertainty that influence the public transport movements impacts the amount of movements by bikes more than that of cars. This means that people are more likely to switch from public transport to the bike than to the car. Lastly, it seems that when the base case is taken, *total\_km* does get almost completely determined by *THUISWERKREDUCTIE*. Taking figure 4.1 in consideration, this means that the effect of the policies on *total\_km* is very different from the effect on *total\_verpl*.

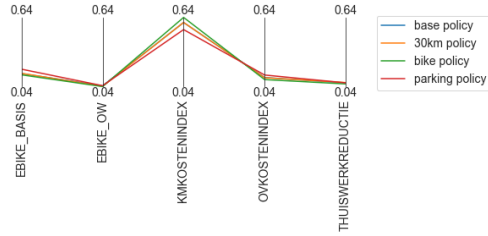
Figure 4.3a, 4.3b and 4.3c show the feature scoring of all uncertainties on *car\_total*, *carshare\_verpl* and *bikeshare\_verpl*, respectively, in the scenarios of every single policy. The complete feature scoring of all the policies is found in appendix B.

The bike policy shows some difference in the feature scores in comparison to the other policies. Especially in the effect of *KMKOSTENINDEX*, *EBIKE\_BASIS* and *EBIKE\_OW* on *carshare\_verpl* and *bikeshare\_verpl*. This means that there is some interaction between the bike policy and these uncertainties. The e-bike related uncertainties seem to have less influence while the car related uncertainties seem to have more influence. A logical explanation is that the fact that using the bike is already made more attractive the effect of the e-bike related uncertainties diminish. Why the car related uncertainty takes this influence over is not clear.

When the parking policy is active there is even more difference. This is especially clear in the effect of *KMKOSTENINDEX* and *THUISWERKREDUCTIE* on *car\_total*. The influence of these uncertainty are traded between each other. What diminishes by *KMKOSTENINDEX*, increases by *THUISWERKREDUCTIE*. The explanation for this is twofold. Firstly both parking costs are *KMKOSTENINDEX* increase the cost of driving a car. The fact that it is already expensive to park in the city the effect of high costs per km diminish. In addition, possibly movements for the working motive are not influenced very much by high parking costs and therefore the reduction in movements due to people working from home stays does not decrease. Due to the nature of feature scoring it is hard to pin down the exact reason behind the behaviour.



(a) A parcoords diagram of the feature scoring of all uncertainties on *car\_total* (b) A parcoords diagram of the feature scoring of all uncertainties on *carshare\_verpl*



(c) A parcoords diagram of the feature scoring of all uncertainties on *bikeshare\_verpl*

Figure 4.3: Parcoord plots of the feature scoring of all uncertainties on three different outcomes

### Feature scoring with changing policies

The last feature scoring looks at the feature scoring when policies are changing in level. In figure 4.4 there are a couple of interesting things. The first is the low effect of *BIKE\_SPEED* on *carshare\_verpl* and the high effect on *carshare\_km*. If this knowledge is combined with the fact that *PT\_KP* has a high effect of *carshare\_verpl* and a slightly lower effect on *carshare\_km*, this means that one of the two policies has a differing effect between the two outcomes.

The second interesting thing is the high effect of *BIKE\_SPEED* on *totaal\_km*, but not on *totaal\_verpl*. Combined with the fact that *THUISWERKREDUCTIE* has the opposite happening, it can mean two things. Firstly, it can mean that people working from home mostly effects the short movements or that more bike use shortens the path of people compared to using another type of vehicle.

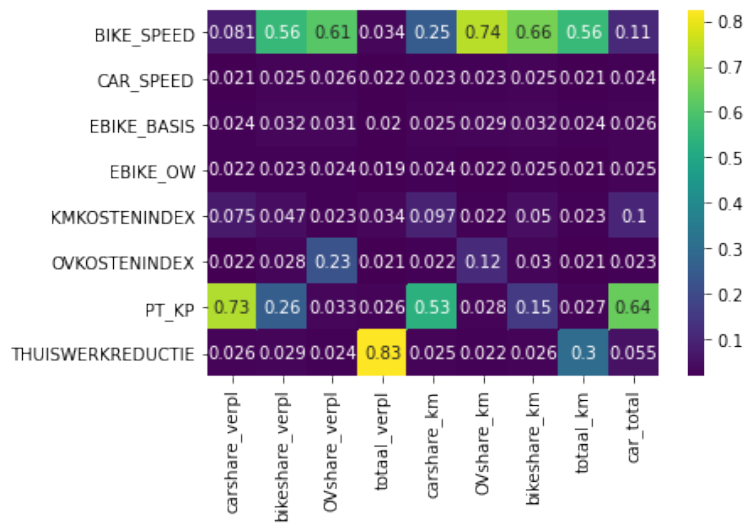


Figure 4.4: Feature scoring with changing policies

The last interesting outcome is that the main kpi *car\_total* is mostly determined by *PT\_KP*. The uncertainty *KMKOSTENINDEX* and the policy *BIKE\_SPEED* also have some impact, but significantly lower than *PT\_KP*. This means that, based on this analyses, parking costs are the main policy to influence the total distance travelled by car.

#### 95% confidence intervals of feature scores

The application of extra-trees can give different feature scores for every run. This is due to the randomness in which the trees are picked. One key determinant of the differences between feature scores is the amount of scenarios that are run. In order to check how much the feature scores differ, 100 feature scores are calculated. These feature scores allow for the calculation of the 95% confidence interval. The 95% confidence interval is calculated for the base case, the case with policies and the case with changing policies. The calculation of the other three policies is found in appendix B. Only the 95% confidence interval of the outcome *car\_total* is taken, due to its importance in this research.

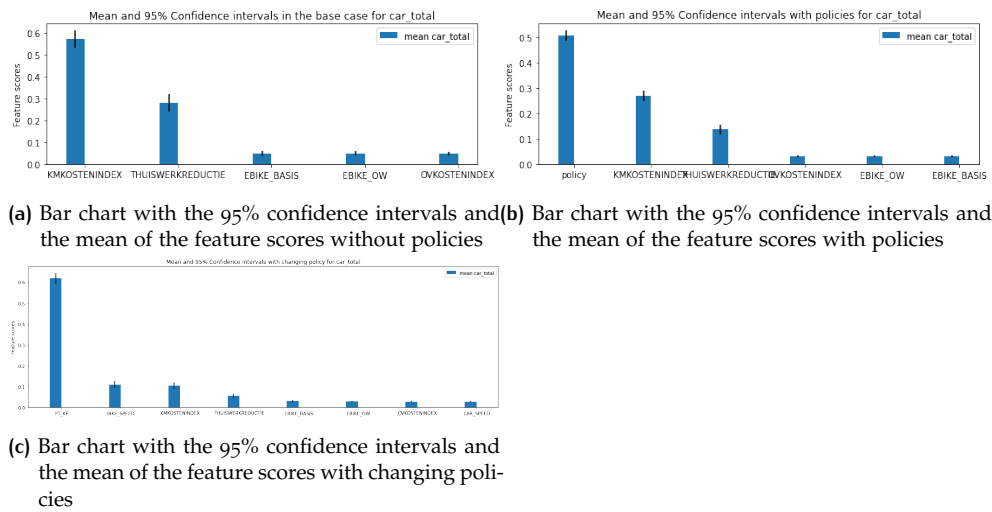


Figure 4.5: Bar chart with the 95% confidence intervals and the mean of three different sets of feature scores

The 95% confidence intervals of the feature scores should not be too large. Since the feature scores are always between 0 and 1 a difference of 0.1 between the upper and lower bound is already quite large. From figure 4.5a it becomes clear that the biggest range is found for both *THUISWERKREDUCTIE* and *KMKOSTENINDEX*. The other two analysis, in figure do 4.5b and 4.5c, show a smaller range for the 95% confidence interval than the one without polices. It should be noticed that the analysis without policies uses only 250 scenarios. Therefore the extra-trees method has less data to work with. The 95% confidence intervals for all three analysis do not raise red flags. The differences between the upper and lower bounds do not change the conclusion of the feature scoring. However, one cannot take the numbers as exact due to the ranges being too large for that. Running more scenarios could get the 95% confidence interval smaller, then the feature scores can be taken as exact more easily. Due to the long run time of the model, the exact amount of runs to achieve this remains unknown.

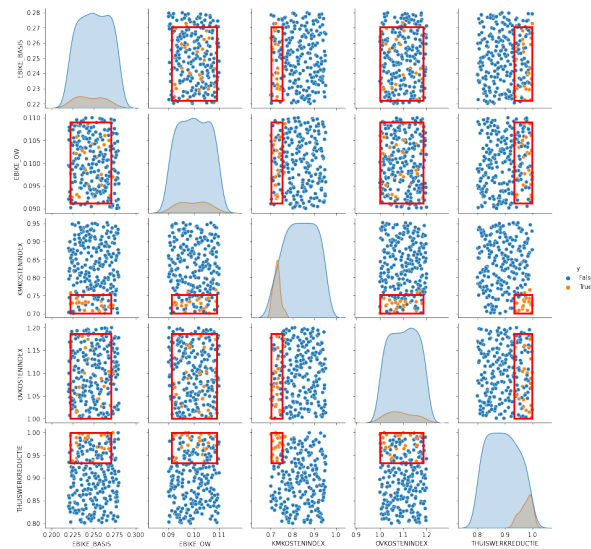
#### 4.2.2 Step 3: Scenario discovery

PRIM is applied to find scenarios of interest. In chapter 2 the method is explained. PRIM is applied for the base case without policies, the case where the policies are applied and the case where the policies change in value. This shows both the effect

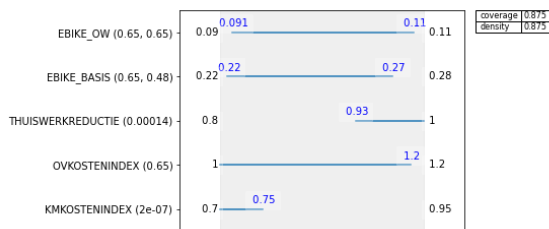
of the uncertainties, the policies and the uncertainties with the policies applied. The PRIM algorithm is applied twice for all three analysis. First for the 10% worst performing scenarios for the amount of distance travelled by car. Second for the 10% best performing policies. The analysis is performed by both providing a graphical overview and a statistical overview. The statistical overview shows the quasi p-values (qp). The most commonly used p-value to show if a variable is statistically significant is 0.05 (Kwakkel, 2019). Thereby, the ranges for variables that determine the box and the coverage and density are shown in these figures.

**Base case - 10% worst performing scenarios**

Figure 4.7a shows the results of only the base case in a graphical manner. Orange points are values that are of interest and the blue ones are not. PRIM tries to find a box within a n-dimensional space which has both high coverage and high density. The coverage is the total amount of orange dots divided by the amount of orange dots within the box. The density is the total amount of dots in the box divided by the dots within the box that are orange. From the qp-values in figure 4.6b, which are shown in brackets behind the variables, it is clear that only *KMKOSTENINDEX* and *THUISWERKREDUCTIE* are statically significant. The upper right corner of figure 4.7a shows a box with a density of 0.875 and a coverage of 0.875, as can be found in figure 4.6b. This means that it does not perfectly explain the data, but still finds a good approximation. The box is made for high values for *THUISWERKREDUCTIE* and low values for *KMKOSTENINDEX*. This is a logical outcome since when more people travel for work and the costs for driving a car are low, the total distance travelled by car increases.



(a) Graphically

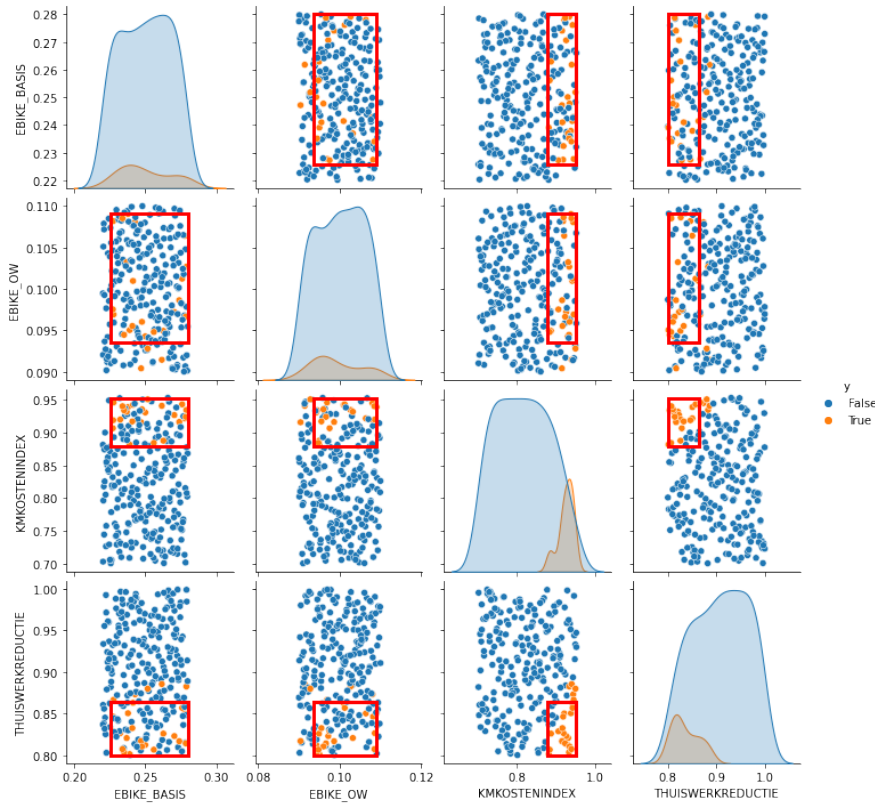


(b) Coverage, denisty and qp values

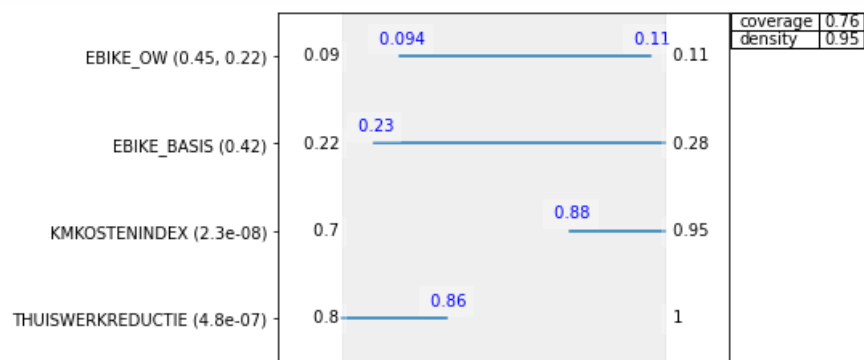
Figure 4.6: PRIM in the base case for the 10% worst performing scenarios

### Base case - 10% best performing scenarios

The results of PRIM in the 10% best case scenarios is almost the exact opposite of PRIM in the 10% worst case scenarios. In figure 4.7 it is clear that the box is created for high values of *KMKOSTENINDEX* and low values for *THUISWERKREDUCTIE*. Figure 4.7b does show a higher density and a lower coverage. This means that for these values almost all cases return true, but there are a significant amount of cases that are true for lower values of these two variables.



(a) Graphically



(b) Coverage, density and qp values

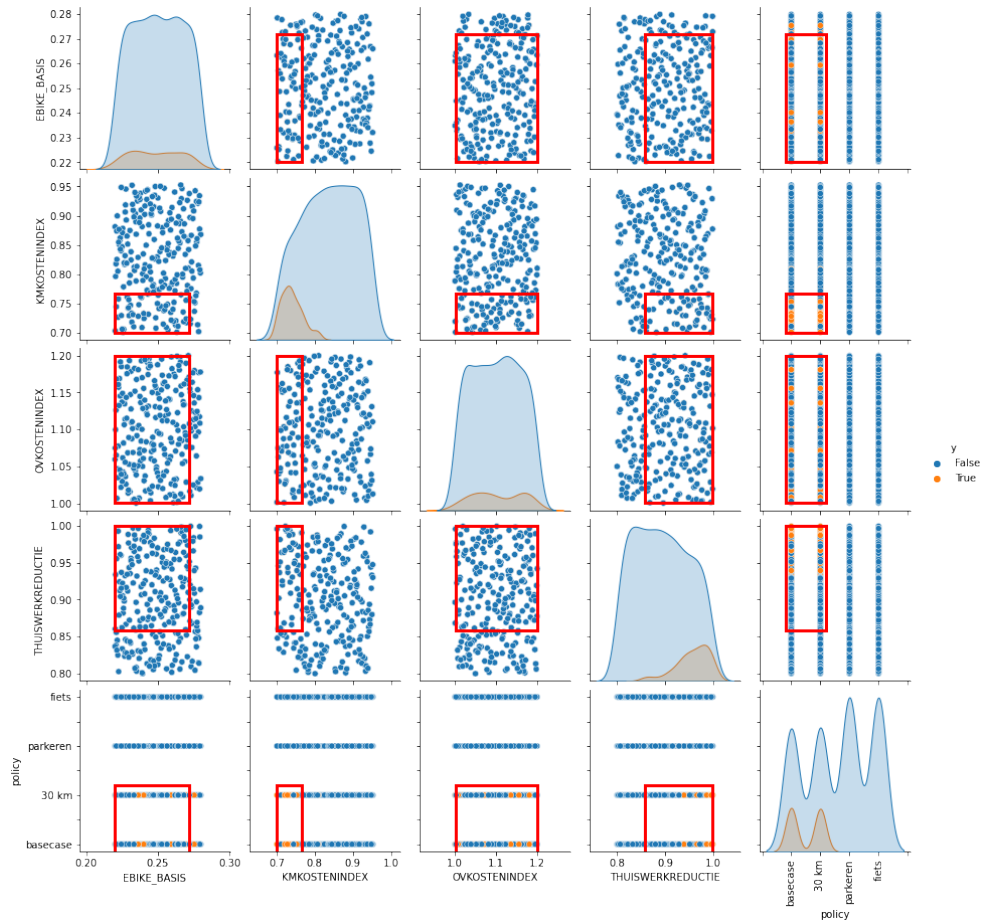
Figure 4.7: PRIM in the base case for the 10% best performing scenarios

### With policies - 10% worst performing scenarios

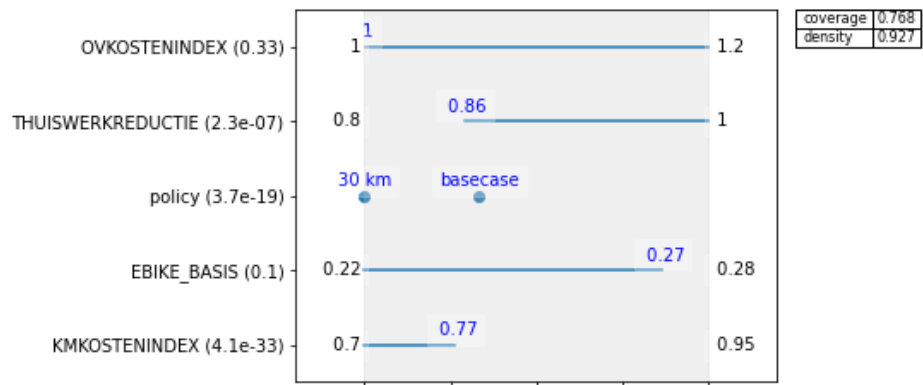
When the policies are also used in the PRIM analysis the outcome is different. Figure 4.8a shows the results graphically. This figure shows that it is impossible to recognize orange dots within the boxes of two uncertainties, due to the high level



of overlap between scenarios. The combination of an uncertainty and the policies gives a clearer view. It shows that all 10% worst cases are situated in the 30 km policy and the base case. *KMKOSTENINDEX* is the main uncertainty that determines if a scenario is true or false. However, from the statistical analysis in figure 4.8b it appears that *THUISWERKREDUCTIE* is also statistically significant. This means that the found box is based on three variables.



(a) Graphically



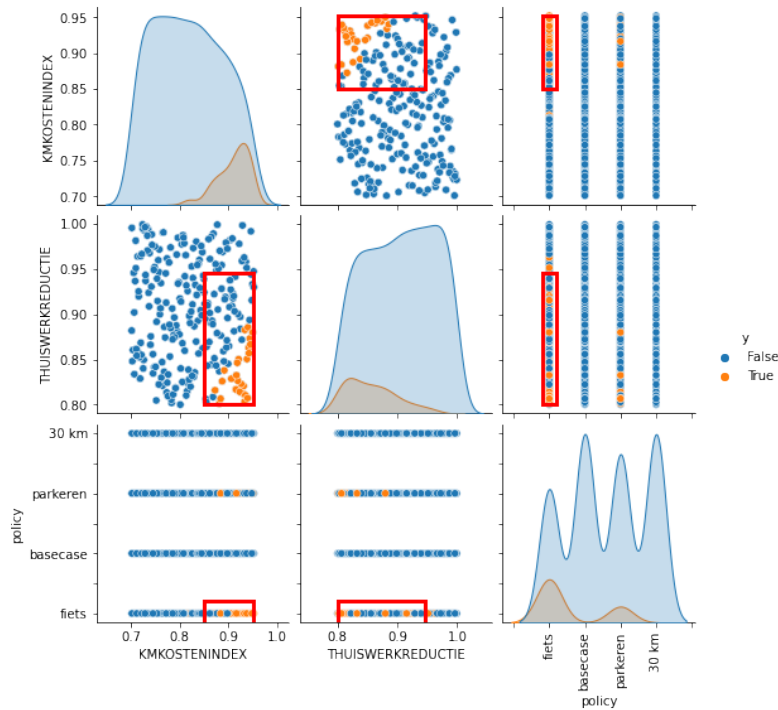
(b) Coverage, density and qp values

Figure 4.8: PRIM for the 10% worst performing scenarios with policies

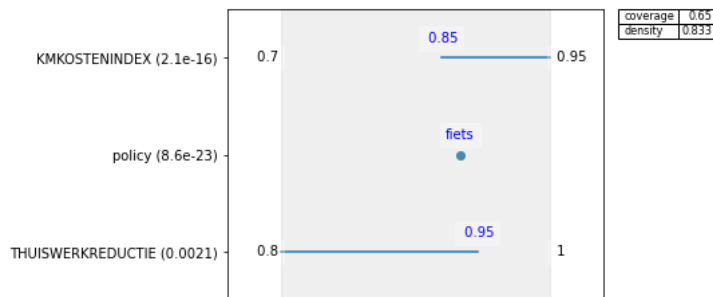
**With policies - 10% best performing scenarios**

Figure 4.9 does show differences with the 10% worst case scenario PRIM analysis. The effect of *KMKOSTENINDEX* and *THUISWERKREDUCTIE* seems to be more or

less the same. The difference is found when looked at the effect of the policies. In this case the fiets policy encapsulates almost all the cases of interest. This means that this policy is very powerful in achieving a good scenario outcome. However, there is quite a low coverage. This is mainly caused by the fact that the parking policy also has some cases of interest. A possible reason for not taking these scenarios in the box is a big decrease in density.



(a) Graphically

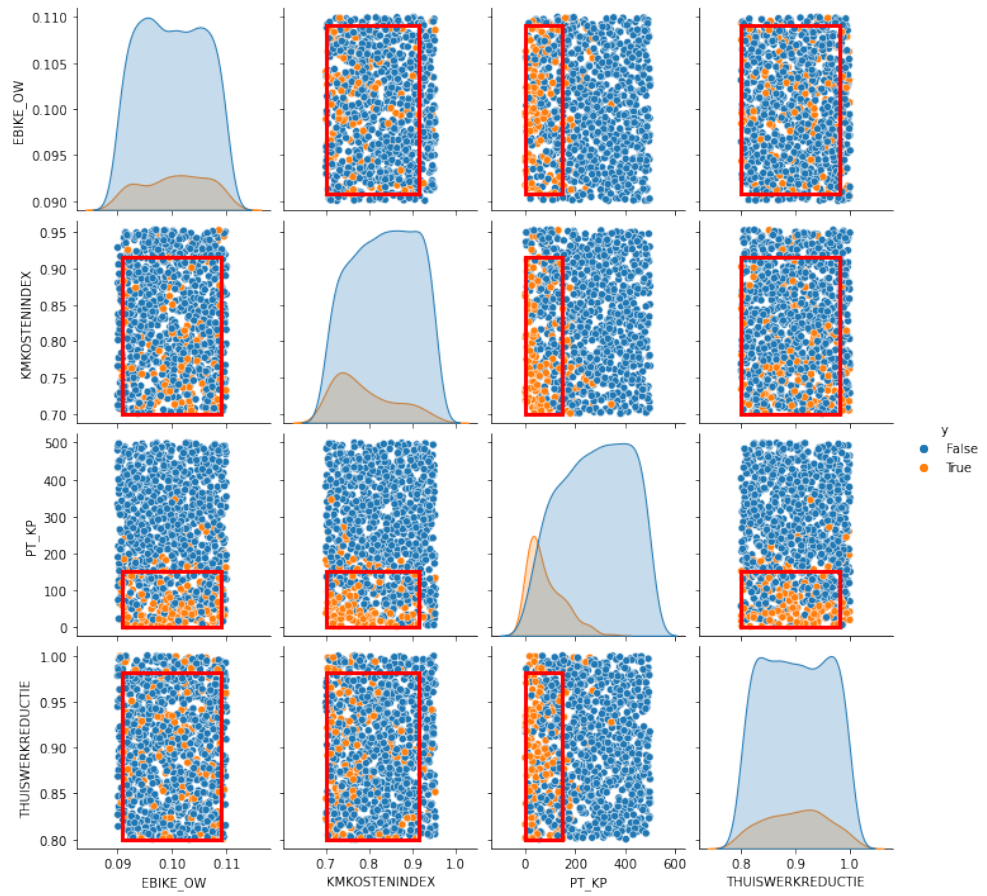


(b) Coverage, density and qp values

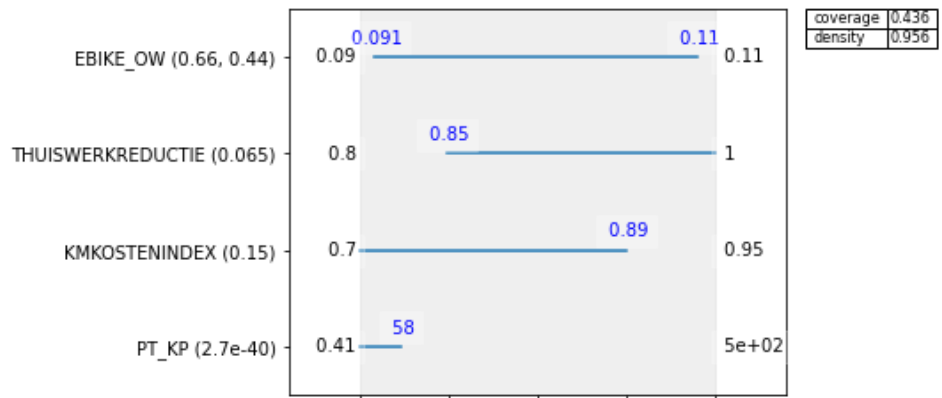
Figure 4.9: PRIM for the 10% best performing scenarios with policies

### With changing policies - 15% worst performing scenarios

Figure 4.10 shows the PRIM analysis of the 15% worst performing scenarios with changing policies. In this case 15% is chosen instead of 10% because PRIM fails to find a solution for 10%. PRIM has difficulty finding a box with both a high density and a high coverage at 15% as well. The coverage is only 0,436 and the density is quite high with 0,956. This means that this box makes up for less than half of all cases of interest. As can be seen in figure 4.10b only a low value for  $PT_{KP}$  is statistically significant. The effect of  $KMKOSTENINDEX$  and  $THUISWERKREDUCTIE$  has diminished so much that it became statistically insignificant. This does explain the reason for not finding a box for only the worst 10%.



(a) Graphically



(b) Coverage, density and qp values

Figure 4.10: PRIM for the 15% worst performing scenarios with changing policies

### With changing policies - 10% best performing scenarios

When looking at the best scenarios with changing policies, PRIM is able to find a 10% best scenarios. Therefore, not 15% but 10% is chosen. Figure 4.11 shows the analysis. The box found is determined by four variables which are all statistically significant. The two uncertainties which failed to be significant in the previous analysis are significant here. The *BIKE\_SPEED* is the second policy alternative that is responsible for the system to perform well. However, there is still a quite low coverage. This means that there are still many cases which are not determined by these four variables. The reason for this is the higher level interaction effects between the policies and variables. This causes the system to perform more variable than in the first four analysis.

## 4.3 STEP 4: ROBUSTNESS ANALYSIS

A robustness analysis can be performed with different robustness metrics. The choice for a certain metric depends on the preferences from the side of the researcher and the decision makers (Kasprzyk et al., 2013). Since the preference of the decision maker is unknown, multiple metrics are used in this research.

**Table 4.1:** Robustness scores, between brackets show if a high or a low value is good

	Minimax regret (min)	90th percentile minimax regret (min)	Maximax (min)	Mean-variance (min)	Starr's domain criterion (max)
Base case	$3,39 \times 10^4$	$3,23 \times 10^4$	$3,00 \times 10^5$	$4,23 \times 10^9$	39
30 km policy	$3,37 \times 10^4$	$3,21 \times 10^4$	$3,00 \times 10^5$	$4,23 \times 10^9$	40
Fiets policy	0,0000	0,0000	$2,70 \times 10^5$	$3,61 \times 10^9$	222
Parking policy	$9,96 \times 10^3$	$8,94 \times 10^4$	$2,78 \times 10^5$	$3,55 \times 10^9$	199

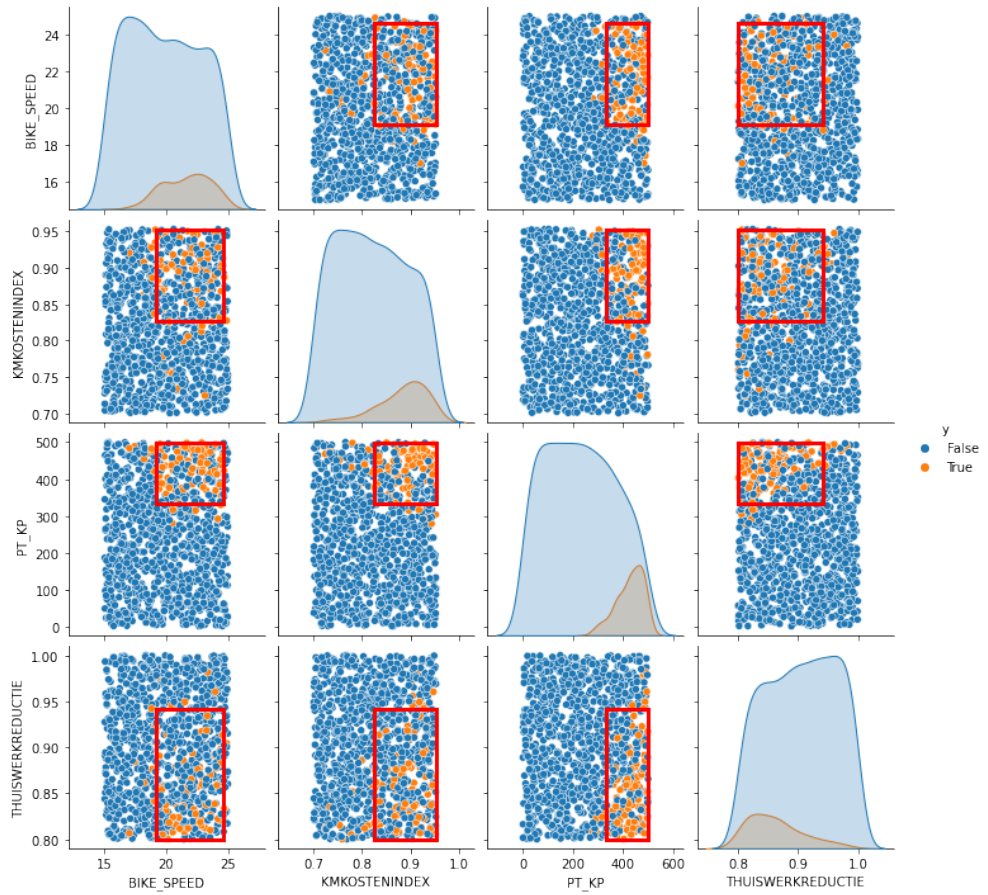
Table 4.1 shows the robustness scores of every policies on every metric used. All the metrics are based on how the policies score on *car\_total*. The best performing policy based on the minimax regret metric is the *Fiets policy*. The minimax regret for this policy is 0. This means that in every scenario that is tested this policy performed best. After the *Fiets policy* the *Parking policy* performs best with the other two policies performing almost equally bad. The 90th percentile metric does not change the outcome that much. This means that the minimax regret metric is not based on outliers in the data.

A feature scoring is performed for only the regret to see how the uncertainties and policies influence the regret. The results of this are found in table 4.2. The scores show the high impact of the policies. Around 91,7% of the regret is determined by the policies. All the uncertainties have an almost equal, and very small, effect on the regret. This means that the difference between outcomes in the same scenarios are greatly influenced by the policies and not that much by interaction effects between policies and uncertainties.

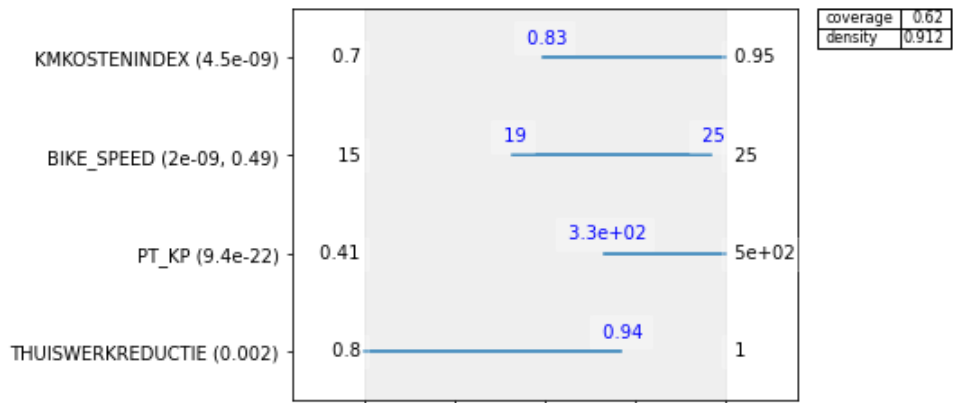
**Table 4.2:** Feature scoring on regret

Variable	Feature score	95% confidence interval
Policy	0,917	0,903 - 0,932
THUISWERKREDUCTIE	0,017	0,013 - 0,021
OVKOSTENINDEX	0,016	0,013 - 0,020
KMKOSTENINDEX	0,017	0,013 - 0,021
EBIKE_OW	0,016	0,012 - 0,020
EBIKE_BASIS	0,016	0,012 - 0,020

The maximax metric shows which policy has performed best in a single scenario. Knowing that the minimax regret is already 0 for the *Fiets policy* it is already de-



(a) Graphically



(b) Coverage, density and qp values

Figure 4.11: PRIM for the 10% best performing scenarios with changing policies

ductible that this policy also has the best maximax score. The *Fiets policy* can achieve a *car\_total* of  $2,7 * 10^5$ . This is a bit lower than the *Parking policy* and around 10% lower than that of the other two policies.

The mean-variance metric shows which policies performs the most consistent while having a low mean value. Now the *Parking policy* performs the best. From the regret metric it became clear that the *Fiets policy* performs the best in every single scenario. Which means that the *Fiets policy* has a lower mean and therefore that the *Parking policy* has a lower variance. It performs more consistent. The other two policies are again the two lowest.

The last robustness metric used is Starr's domain criterion. The criterion is set on a *car\_total* of  $3,1 * 10^5$ . Which means that scenarios that score above  $3,1 * 10^5$  are not adding to the score and the other way around they are adding to the score. This value is based on the median value of the total data set. This computation shows that *Fiets policy* almost always performs better than the median value. The *Parking policy* performs in around 20% of the cases worse than the median. Logically, this means that the other two policies have some policies that perform better than the median. If the municipality of Groningen would use the the median value, the *Base case* and the *30 km policy* could perform good enough in around 15% of the scenarios. In that case they could consider keeping the *Base case* if other considerations, like budget, play a big role.

## 4.4 CONCLUSION

The first analysis that is performed is a global sensitivity analysis. This method is mainly used to see which uncertainties and policy are the most influential on the system. This analysis shows that, based on only the uncertainties, the main KPI's are mostly influenced by the cost per km of car use and the amount of people working from home. Thereby, it seems that based on the high influence of the cost per car km on both the amount of movements by car as by bike, that people tend to change from car to bike and vice versa instead of changing to public transport. When the policies are taken into the analysis the underlying effects become more clear. The effects of uncertainties seem to diminish when the policies also influence the outcome which the uncertainty influenced greatly. The last part of the analysis changes the level of the policies. The outcomes for the feature score show that the costs for parking and the speed of bikes are the most influential factors to reduce distance travelled by car. The effect of the uncertainties also seem to diminish by quite a lot. Only on the total amount of movements, the reduction is travels due to people working from home has by far the most impact. The 95% confidence intervals calculated for the feature scores did not show to large ranges. This means that the amount of runs are sufficient for the goal of this research.

The second analysis is the scenario discovery with PRIM. The most impacting uncertainties from this analysis is again the cost per car km and the amount of people working from home. No other uncertainty has a really big impact. The amount of e-bike owners does have some impact, but is not statistically significant in any analysis. With the policies included there are two policies, base case and 30 km, that cause all the 10% worst scenarios and one policy, bike policy, that causes the top 10% best scenarios. So purely base on this PRIM analysis the bike policy is the best policy to achieve the goal of the municipality. However, the cost per car km and the amount of people working from home do influence this solution. If the cost per car km are low the bike policy does not perform in the top 10% and the same goes for a low amount of people working from home. Vice versa this is also true. If the cost per car km or the amount of people working from home are very high, the 30 km policy and the base case wont perform in the 10% worst scenarios. The last scenario discovery analysis is with changing policies. For the 15% worst scenarios the main influence is the low cost for parking. Which means

that low parking costs are not a good solution when low car share needs to be reached. In the best 10% analysis a high speed of bikes and high costs for parking almost always gets a positive outcome. However, low cost per car km and a high amount of people working from home can still effect these solutions to perform badly. The other uncertainties again seem to have very little effect.

The last analysis is the robustness analysis. The robustness analysis showed for 4 of the 5 metrics that the bike policy performed best, the parking policy came in second while the other two performed evenly bad. The bike policy scored 0 on the minimax regret metric. This means that in no scenario another policy performs better. The main cause of this is that there are very little interactions between the uncertainties and the policies. There is a little enhancing effect when for example the bike policy is active and more people own an e-bike. However, this effect seems to be so small that it can not really compensate for the positive or negative effect that a policy has on the system. The feature scoring on regret confirms that the interaction effects between policies and uncertainties are very small. Another remarkable thing is the tiny effect of the 30 km policy. Apparently this policy effects this system by so little that it on average is good for a 0.01% change in car share of total distance travelled. The satisficing metric of Starr does sketch a different picture. Here it shows that the two worst policies do not always perform below what is needed. Therefore, if the budget is small, doing nothing can still prove useful.

# 5 | DISCUSSION

In this research a deep uncertainty analysis is performed on a macro-level traffic model of a big city in the Netherlands. Firstly, this chapter discusses how the results should be interpreted and if they can be generalized to other macro-level traffic models. Secondly, the implication of the results is discussed. Thirdly some limitations of the research are given. Finally, the possibilities for future work are discussed.

## 5.1 INTERPRETATION

This research focuses on the application of the RDM method on a macro-level traffic model of a big city in the Netherlands. By applying the RDM method robust decisions, which take deep uncertainty into account, can be taken. The results show that every step of the RDM cycle can be performed in such a model. In the first step two possibilities arise. On the one hand, the modeller can choose to pick a set of unchanging policies. However, due to the high level of detail in macro-level traffic models the results of such an analysis can get very one sided. This is obvious in the fact that the bike policy performs best in every single scenario. On the other hand, the modeller can choose to change the value of the policies. This way a combination of viable policies can be found. Nevertheless, the run times of macro-level traffic models are too high to fully exploit the usefulness of this type of analysis. With changing policies the space in which is sampled becomes larger and therefore needs a lot more scenarios to cover the sample space. Another possibility that changing policies provide is optimization. However, this does require lower run times as well.

In the second step of the RDM cycle, where the model is run over a wide set of scenarios, has two important considerations. The first is how to connect the traffic model to the EMA\_workbench. In the case presented in this research PTV Visum is used. This has a built-in python console which makes the connection easy to implement via a COM interface. Other software written by PTV, like PTV Vissim, have these built-in python consoles as well. Different traffic model software might have a harder time connecting to the EMA\_workbench. However, the EMA\_workbench is not the only way to run a wide set of scenarios. For example, PTV Visum could perform the same tasks via reading a dataset to determine scenarios. This does require some extra coding within the traffic model. The second consideration is how many scenarios are run. Since the run time of macro-level traffic models is quite long, it is very dependable on the amount of uncertainties and levers and the available time of the modeller. This research shows that for 5 uncertainties and 4 fixed policies, 1000 scenarios are sufficient to get a meaningful analysis. However, this does not mean that this is sufficient for other traffic models. It is difficult to pinpoint the exact amount of scenarios that have to be run to do a meaningful analysis. When there is a lot of interaction between variables 1000 might be too little and when there is no interaction between variables, less runs could suffice. Via an iterative process the modeller should be able to pinpoint the right amount of runs for the problem at hand.

The third step shows the effectiveness of both GSA and scenario discovery in macro-level traffic models. GSA reveals how the outcomes are influenced by both



the uncertainties and the policies. This helps the decision-maker to see which uncertainties are important to track and which policies can be helpful in achieving a goal. The case of this research did show that macro-level traffic models have low interaction effect. Between different policies the effect of uncertainties on the outcomes changes very little. This effect can be assigned to over fitting, which is described in more detail in sub-chapter 5.3. It can be assumed that this is the case for other macro-level traffic models in the Netherlands, since most of these models are based on the same base model as the Groningen model, namely the LMS and NRM (Rijkswaterstaat, 2021). The research of De Jong et al. (2007) also looked at the LMS since this is so widely used in the Netherlands. The scenario discovery part shows that PRIM is able to find scenarios of interest in the sample space. Based on preferences from the side of the decision maker a threshold value, which has to be reached, can be set. With the help on the EMA\_workbench it should be easy to conduct such an analysis.

The fourth step of the RDM cycle consists of a trade-off analysis. This trade-off analysis is based on robustness metrics. In this research multiple metrics are used to show the possible differences in ranking. The minimax regret metric makes it clear how low the interaction effects are in macro-level traffic models. The outcomes of the regret values are almost exactly the same between different policies for the same scenarios. However, there are some interaction effects visible, because one policy shows a better score for the mean-variance metric, while performing worse in every scenario. This means that it does matter which metric is chosen to use in the scoring of policies in a macro-level traffic models. Metrics that compare the performance between policies within the same scenario and satisficing metrics do not change insight into the results. These metrics put forward the exact same policy as the best due to that policy performing best in every single scenario. Metrics that look at consistency, like mean-variance, can show a different perspective than the previous mentioned metrics. A combination of a metric that measures consistency and a regret/satisficing metric would hand the decision maker with the most useful information.

In conclusion, the RDM cycle is applicable to macro-level traffic models of large cities and can prove to be very helpful in coming to more robust decisions. However, there are some characteristics of traffic models which make the execution of the steps more difficult. Accounting for these characteristics is an obstacle that can be overcome with the tools provided in this section.

## 5.2 IMPLICATIONS

As mentioned in the introduction (1) there were already some researches that addressed uncertainty in traffic modelling. Especially the researches of De Jong et al. (2007) and Petrik et al. (2018) tried to cope with the uncertainty. The difference with these researches is the size and the methods use. Petrik et al. (2018) looked at 100 uncertainties over 500 scenarios. The problem with this is that just a small part of the uncertainty space is uncovered. With 100 uncertainties even 100000 runs have trouble with searching the whole uncertainty space. The eventual analysis that is performed, is a look at the mean, standard deviation and the standard deviation divided by the mean, which they call the coefficient of variation. This research shows that more in-depth analysis, like PRIM and feature scoring, are possible to perform on the scenarios. The research of De Jong et al. (2007) did look at a similar set of uncertainties, policies and outcomes. However, they only sampled 100 scenarios of which 50 had a policy and 50 had no policy. The research in this thesis does a more extensive analysis and tries to represent the uncertainty space more broadly. This is done by running a 1000 scenarios for only 5 different uncertainties. This proves that it is certainly possible to apply deep uncertainty analysis in traffic models. Thereby, it proves useful in finding a robust decision and the sensitivity of the system to mul-

multiple uncertainties. It allows for scenario discovery techniques and global sensitivity analysis.

The analysis could have applied more complex methods. [Bonham et al. \(2020\)](#) used the Multi Objective Robust Decision Making (MORDM) method. In this method the objective first gets optimized to then be tested over a set of future scenarios. This method is quite similar to the RDM method, with the exception that it optimizes the policies instead of just test some chosen policies. Optimization is a computational expensive procedure. Therefore, it is difficult to do this for traffic models due to the long run times. In other words, deep uncertainty research have produced more advanced techniques, but this research has yet to prove that this is also possible for traffic models. Micro level traffic models could be more suited for such research. However, this has to be tested in future researches.

This research has proven that RDM can be applied to macro-level traffic models. [Groves et al. \(2019\)](#) proved that RDM was applicable in multiple sectors. They conducted the analysis for water resource management policies and climate change mitigation's policies. Now, traffic management policies can be added to the list of sectors in which the the RDM can be usefully applied. For more advanced techniques, more research is needed.

### 5.3 LIMITATIONS

Every research has some limitations that have to be taken into account when looking at the results and conclusions. This research has three main limitations, which are discussed below.

#### *Run time and time frame*

One of the biggest limitations of this research is the run time of the Visum model. Even though the run time has been lowered from 22 hours and 45 minutes to 12 minutes per run. This combined with the fact that the time for this thesis was only 4 - 5 months, the quality of the research suffered. The first thing that makes this apparent in the non-circular way of conducting the RDM cycle. As can be seen in chapter 2, the idea with RDM is to test some policies on a system and than adjust the system and come up with new policies. Due to the high run time, adjusting the model and choosing new strategies was hard to fit in the given time frame. The long run time and short time frame, also made it hard to apply any more advance techniques. The EMA\_workbench allows for optimization of the solution and optimization of the robustness. However, running such techniques would take such a long time that it could not be performed in this research. Note that this is not necessary for RDM, but it is necessary for many other deep uncertainty methods.

The run time has caused another limitation. That is the amount of procedures that can be run inside the Visum model. The way the run time has been decreased is by reducing the amount of procedures. Therefore, only the direct tours are included and there are not multiple iterations conducted for the model. However, as shown in chapter 3, only taking the direct tours does not change the outcomes of the model in a way that is hindering. The fact that only one iteration is run reduces the reliability of the model run. Nevertheless, the high amount of single scenarios run do already account for the randomness of the model.

Since macro-level traffic models in the Netherlands are all based on the LMS and NRM models, the same problems occur for the same type of models. It is an inherent problem of macro-level traffic models due to the way the algorithms work and the high level of detail in the models. Sub-section 5.4 elaborates more on ways to deal with these run times.

### *Over fitting*

In the model there are very little interactions effects between the uncertainties and the policies. A very likely reason for this is that the model has been over fitted (Hawkins, 2004). In the pursued of a model that has a very great level of detail which describes every possible relation, single variables lose their effectiveness. It is debatable if this is a limitation for the purpose Sweco uses it for. However, it is a limitation for the purpose of this research. It makes the robustness analysis very one sided. In every scenario all the policies ranked in the same order and only one metric slightly preferred another policy. When the policies change in level there are more interactions effects visible, but doing an optimisation for this is to time-consuming due to the previous mentioned limitation. The value of the analysis does decrease a bit due to the little interaction effect. You could argue that, in this case, to get the most robust solution you could run it with any set of uncertainties since it always ranks the policies in the same order. However, especially Starr's domain criterion does show its value here. If there is a criterion that should be reached this metric can still show if a policy does suffice. The little interaction effect also effects the other two analyses. As shown, the comparison of feature scores between policies could not show the interactions between policies and uncertainties. While it is expected that, for examples, making biking more attractive and having more e-bike users would show some form of enhancing effects.

### *Limited variable set*

The amount of uncertainties, policies and outcomes used in the analysis is relatively low, 1 kpi of interest, 4 policies and 5 uncertainties. This is mostly caused by the fact that the Visum model was already constructed. Adding new uncertainties and policies could disrupt the complex working of the model. Thereby, the outcomes are to be taken from what is generated by the model. Outcomes like costs or expected congestion would add a lot of value in the analysis. Uncertainties like the increase in e-bike use for educational purposes are of very little influence to the model. Because of the fact that there are some low impact uncertainties, it would be interesting to add some uncertainties on for example demographic data. The same goes for the policies. The 30 km policy had almost no effect on the outcomes. Trying more different policies would add value to the total analysis. To conclude, with a broader set of variables, the analyses could have been more in-depth.

Regarding the outcomes it is debatable if it really limits the research. As mentioned in chapter 1 an often used metric is the distanced travelled for a certain mode. In many researches such an outcome is the only outcome. However, this characteristic of traffic research does limit the possibilities. For example, with the addition of costs and expected emissions a trade-off can be made between these conflicting outcomes. The Visum model in this research did not have any costs involved in the model.

### *Conclusion limitations*

The limitations and obstacles found in this case study are expected to be found in other macro-level traffic models. Especially in these type of models for large cities in the Netherlands. As mentioned, most Dutch traffic models are built from the LMS and NRM models. Therefore, run time limitations will be the same, the models are over fitted and the available set of variables is limited. The next sub-chapter goes deeper into possibilities to deal with these limitations.

## 5.4 FUTURE WORK

This section focuses on possibilities for future research. Some parts are written from the point-of-view of the case which is run for this research. However, since macro-level traffic models are very similar, especially for cities in the Netherlands, all proposed future research can be conducted on similar cases. It is debatable if the same is true for traffic models from other countries. For these models it should first be checked if the models share the same characteristics as that of macro-level traffic models of cities in the Netherlands.

Despite some limitations to the current conducted research, it does open up some roads to new or expanding research in this topic. The first possibility is to built on the existing work with none changing policies. As mentioned in the limitations, the variable set limits the research. Adding for examples costs of a policy as an output or the amount of car-owners as an uncertainty would broaden the scope of the research. However, it is hard to add new variables to the model without messing up its structure. Therefore, to achieve this expert knowledge from the developer of the model is needed. From the outcomes side, executing this on any case opens doorways to a trade-off analysis with conflicting goals. In this research there were no conflicting goals.

The second possibility is to go further with the changing policies. The changing policies have one big advantage, they can be optimised on both outcomes and robustness. This allows for more advanced techniques like Multi Objective Robust Decision Making (MORDM) or Dynamic Adaptive Policies (DAP) (Bartholomew and Kwakkel, 2020). Before this is possible a couple of things are required. First of all, there should be more research in possible possibilities and the realistic values these policies can take on. Second of all the run time has to be lowered drastically. There are several ways to achieve this. Reducing the size of the existing model is possible, but is very hard without changing the model quality. The biggest time consumer is in the iterations of the computation of the 2200 by 2200 matrix, which equals 5 million computations. Reducing this matrix to a 100 by 100 would already reduce the amount of computations by a factor of 500. Visum does allow distributed computations. Coupling more calculating computers to the model could reduce run time quite a lot. However, this does require multiple node to run on and a wide range of different licenses. The first problem can be solved by a provider like Amazon Web Services (AWS). The second problem should be discussed with the provider of the license PTV. There are definitely possibilities to construct something like this to reduce total run time, but it could get costly very quickly. A probably more time consuming and realistic approach would be to remake the model, but way less detailed. This way the most optimal and robust policies can be found in a smaller model and these policies can than, for more extensive calculation, be applied in a bigger model. This type of modelling is also referred to as multi-resolution modelling, as in the paper of Yang et al. (2012). The biggest challenge according to that paper is keeping the different models consistent. This is definitely the biggest challenge in traffic models as well, due to the high level of detail in the bigger model.

The third possibility is to do a deep uncertainty analysis on a not yet existing model. If before the model is made the decision is made to do a deep uncertainty analysis, the model can be adjusted to the needs of a deep uncertainty analysis. In combination with the previous idea, two different models can be made simultaneously. The first is quick and basic to do the deep uncertainty analysis and the second is slow and very detailed to do more route assignment based calculations. The idea is that knowing the purpose beforehand can make an integration between two such models easier. This reduces the severity of the consistency problem of multi-resolution modelling.

Lastly, one could argue if there really is a need for such a detailed model. The "Mobiliteitsvisie Groningen 2040" is a very broad prediction of the year 2040. With the time span of 20 years their ought to be a lot of deep uncertain factors. For ex-

ample the prediction for demographic data is now taken as a known factor, while it is impossible to predict such data perfectly. If the model is so uncertain overall, this also means that outcomes like congestion on a street falls victim to this uncertainty. The model tries to predict those outcomes very precisely with a lot of heavy algorithms while the future always turns out differently. So there seems to be a trade-off between preciseness and taken deep uncertainty into account via, for example, RDM. In most traffic researches preciseness is valued higher than doing an uncertainty analysis. The question rises, how close do these precise future predictions come to real future? If these predictions are far off, using a deep uncertainty analysis method might proof more valuable than that exact prediction. Thereby, the analysis in this research has shown that due to over fitting the model fails in recognizing small changes. All in all, these reasons show that it is definitely worth to take a look at less detailed models.

# 6 | CONCLUSION

The final chapter of this thesis starts with the answering of the sub-questions and the main research question. After that some recommendations are given to the municipality of Groningen.

## 6.1 ANSWER RESEARCH QUESTION

In the introduction (1) a set of 4 sub-questions is given. These sub-questions are listed below with an answer and short discussion.

- *Which method can be used to take deep uncertainty into account in a traffic model?*

The answer to this question is found in chapter 2. There are multiple methods that try to deal with uncertainty in order to improve decision making. Robust Decision Making (RDM) is the method used in this thesis. RDM improves decision making on multiple facets. It allows the decision maker through an iterative process to improve the way the system is modelled and to find a policy that performs most robust in a wide set of plausible futures. A robustness metric is chosen by the decision maker to let their preferences shine through. This preference is based on level of risk aversion and if the solution should satisfy or should have the lowest possible regret. The difference with decision making without regarding deep uncertainty is that scenarios which are normally not taken into account are taken into account. This can filter out possible combinations within the uncertainty space that would make a policy fail.

- *What are the important deep uncertainties, policies and outcomes in traffic models?*

Chapter 3 answers this question. There are initially 5 uncertainties which the Visum model runs experiments for. These are the amount of people owning an e-bike from normal to educational purposes, reduction in home-work movements due to people working from home and the cost for using the car and the public transport. The original upper and lower bound for these uncertainties values is taken from the WLO scenarios. For the deep uncertainty analysis these values are partly taken as the same and partly changed due to previous optimistic/pessimistic predictions.

The policies are used in two ways. The first way is by taking a set of 4 policies that cannot change. One policy lowers the speed on some roads from 50 to 30, one policy changes some roads so only bikes can ride on them at a higher speed, one policy that increases the size of the area where parking costs are charged and lastly a policy in which the municipality does nothing. More policies can be chosen, but these four seem the most distinctive. The second way is by changing the level of the policies. So the same policies are all implemented at the same time, but, for example, parking costs can differ between 0 and 5 euros. This can show what the correlations between policies and outcomes are.

There are in total 9 outcomes that are researched. However, to the municipality of Groningen the distance travelled by car within the city is the most important outcome. Literature in chapter 1 shows that this is a commonly used outcome for traffic research. The other outcomes help in understanding why a reduction in distance travelled by car takes place.

- *Which uncertainties and policies is the model sensitive to?*

This question is answered in the scenario discovery part of chapter 4. The scenario regions that are looked for are the 10% best and 10% worst case scenarios. Regarding the uncertainties, in both cases the amount of people working from home and the cost per car km have the most influence. When policies are included the 10% best performing scenarios are always found when the bike policy is active. The 10% worst performing scenario are found by both the 30 km policy and the base case. If the policies are active there is a clear reduction in the effect of the uncertainties. Especially of people working from home. This is the case because the effect of the policies is higher than that of the uncertainties. The scenario discovery is also done with changing policies. To achieve a scenario that performs in the worst 10%, having low parking costs is the most effective policy. To achieve a scenario that performs in the best 10%, more variables are of interest. In this case having high parking cost, a high bike speed and high cost per km for car has the highest change to perform well.

- *Which policies perform best across a wide set of future scenarios?*

In order to answer this sub-question it first needs to be determined when a policy is "best". From the method chosen in the first sub-question the best policy is the one that is most robust. Robustness can be measured in multiple ways. In this research 5 different metrics are chosen which can all have different outcomes, since they all reward other aspects of a policy. Combined with the answer of the second sub-question, the metrics are calculated on the performance of a policy on the total amount of distance travelled by car. In chapter 4 the outcome for all policies on all robustness metrics is found. The bike policy performs the best over almost all metrics. The only metric in which it gets outperformed by the parking policy is the mean-variance metric. This metric rewards a good average value with a low variance. So the best performing policy can not unambiguously be determined. Even though the bike policy has the lowest distance travelled by car in every single future, the parking policy is more consistent over those futures.

Finally, the main research question can be answered.

*"How to account for deep uncertainty in macro-level traffic models to achieve more robust decision alternatives?"*

This research has shown that by applying the RDM method, the decision maker can get more robust decision alternatives. The method is based on running policies of interest over a wide set of futures, which are combinations of values of deep uncertain variables. The complete set of results over these futures allows for the computation of the robustness. However, there is a choose for the modeller in the way robustness is measured. There is a wide list of robustness metrics that differ in what quality of a policy they reward. The choose of metric can change the ranking of different decision alternatives. There are some limitations in detailed traffic models that make it more difficult to apply the RDM method. The main limitation is the run time of the models. For example, the model used for this research initially took 22 hours and 45 minutes to perform one run. The run time had to be lowered to approximately 12 minutes to perform a acceptable amount of runs. This reduced the amount of output that the model could generate, which means that the original analysis performed on the model are not possible to perform on the adjusted model. To conclude, RDM is a good method to account for deep uncertainty in traffic models, but the run time has to be lowered in order to do more meaningful analysis. The chapter 5 goes into greater detail about solutions for the limitations of the model.

## 6.2 RECOMMENDATIONS MUNICIPALITY OF GRONINGEN

The municipality of Groningen has to make a decision on which policies they will apply. Based on this research, changing lanes into fast bike lanes seems, in most cases, to be the most robust policy. In all scenarios, this policy performed better than the other three policies that are tested. However, if the municipality values consistency very high, increasing the parking zones is the better policy to apply. Nevertheless, the results from this research are somewhat narrow. There are more policies that could be tested and there are no combination of policies tested in the model. Therefore it is too easy to choose a certain policy only based on this research.

Outside of the scope of this research there are many ways the municipality can build upon this research. This first interesting way is by testing policies, that the municipality are seriously considering, on their robustness. The way the connection is made between PTV Visum and Python allows for a very easy implementation of other policies. It can open up some areas of interest in the uncertainty space for a policy which could have been overlooked without this analysis.

A second option is to broaden the current model. The current model has a small amount of uncertainties and policies. If there are other uncertainties and policies that the municipality deem important they can be added. A direction of making the run time of the model shorter can also open up a wide set of new opportunities. This could allow for the application of more advanced techniques like DAP or MORDM. This is most interesting with the changing policies. For example, the changing policies can be optimised and over a set of optimal solutions a robustness analysis can be performed. The run time does have to be a lot shorter than is. Finding optimal solution can already take multiple weeks with the current run time. If 50 solutions are found and these are run over 1000 scenarios, the total run time goes into the months. The main place of time reduction can be found in the amount of zones. The mode choice consists out of matrix calculations. The size of the matrices is equal to the amount of zones in the model. This means that the matrices now have a size of  $2200 * 2200$ . If this is reduced to  $500 * 500$  the run time reduction is enormous. The run time problem can also be reduced by using nodes via a provider like AWS. However, this can be very expensive and the fact that PTV Visum runs with licenses has to be considered.

The last recommendation for this case is to make a second smaller model as mentioned in chapter 5. The idea with such a model is that it replicates the behaviour of the bigger more complex model. It should produce close to the same outputs. This model does not have to be made in PTV Visum. There are many other simulation tools that can produce similar outcomes. When choosing the package there are two considerations. One, is it able to produce the same outputs and two, can it produce the same outputs with a shorter run time. When there is a model that fulfills the requirements, the policies that performs best in the smaller model can be tested in the more complex model. This should be an iterative process. First get robust and optimal policies from the smaller model and then test them in the bigger model. If these do not seem to suffice, adjust the smaller model and find new robust optimal policies.

Beside the currently discussed case, there are also possibilities for future cases or other current cases. For future cases there is one big advantage, the model has not been made yet. This means that a bigger and smaller model can be made in parallel. This way the characteristics of the models can be aligned. In addition, the outcomes produced by the models and the uncertainties used are easier to implement. If the municipality is considering a deep uncertainty analysis, this is the best approach. In every research a form of a deep uncertainty analysis can add value. It is up to the municipality to weigh the extra costs of such an analysis to the potential value it can offer. The fact that the connection between PTV Visum and the EMA.workbench is already constructed, makes the cost for applying RDM lower. This should make the consideration of applying RDM easier.





## BIBLIOGRAPHY

- Agop, M., Gavriluț, A., Crumpei, G., and Doroftei, B. (2014). Informational non-differentiable entropy and uncertainty relations in complex systems. *Entropy*, 16(11):6042–6058.
- Armoogum, J. (2003). Measuring the impact of uncertainty in travel demand modelling with a demographic approach. In *Proceedings of the European transport conference (ETC) 2003 held 8-10 October 2003, Strasbourg, France*.
- Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations research*, 41(3):435–449.
- Barceló, J. (2010). Models, traffic models, simulation, and traffic simulation. In *Fundamentals of traffic simulation*, pages 1–62. Springer.
- Bartholomew, E. and Kwakkel, J. H. (2020). On considering robustness in the search phase of robust decision making: a comparison of many-objective robust decision making, multi-scenario many-objective robust decision making, and many objective robust optimization. *Environmental Modelling & Software*, 127:104699.
- Bonham, N., Kasprzyk, J., and Zagana, E. (2020). Robust robustness: A sensitivity analysis of mordm with competing assumptions about future states of the world. In *AGU Fall Meeting Abstracts*, volume 2020, pages H173–07.
- Calvert, S., Taale, H., Snelder, M., and Hoogendoorn, S. (2018). Improving traffic management through consideration of uncertainty and stochastics in traffic flow. *Case Studies on Transport Policy*, 6(1):81–93.
- Christopher Frey, H. and Patil, S. R. (2002). Identification and review of sensitivity analysis methods. *Risk analysis*, 22(3):553–578.
- Curtis, C., McLeod, S., Hultén, J., Pettersson-Lofstedt, F., Paulsson, A., and Hede-gaard Sørensen, C. (2020). Knowledge for policy-making in times of uncertainty: the case of autonomous vehicle model results. *Transport Reviews*, pages 1–21.
- De Jong, G., Daly, A., Pieters, M., Miller, S., Plasmeijer, R., and Hofman, F. (2007). Uncertainty in traffic forecasts: literature review and new results for the netherlands. *Transportation*, 34(4):375–395.
- Douglas, M. J., Watkins, S. J., Gorman, D. R., and Higgins, M. (2011). Are cars the new tobacco? *Journal of Public Health*, 33(2):160–169.
- Flyvbjerg, B., Bruzelius, N., Rothengatter, W., et al. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge University Press.
- Friedman, J. H. and Fisher, N. I. (1999). Bump hunting in high-dimensional data. *Statistics and Computing*, 9(2):123–143.
- Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine learning*, 63(1):3–42.
- Groves, D. G., Molina-Perez, E., Bloom, E., and Fischbach, J. R. (2019). Robust decision making (rdm): application to water planning and climate policy. In *Decision Making under Deep Uncertainty*, pages 135–163. Springer, Cham.

- Hall, J. W., Lempert, R. J., Keller, K., Hackbarth, A., Mijere, C., and McInerney, D. J. (2012). Robust climate policies under uncertainty: A comparison of robust decision making and info-gap methods. *Risk Analysis: An International Journal*, 32(10):1657–1672.
- Hamarat, C., Kwakkel, J. H., and Pruyt, E. (2013). Adaptive robust design under deep uncertainty. *Technological Forecasting and Social Change*, 80(3):408–418.
- Hawkins, D. M. (2004). The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1):1–12.
- Helgeson, C. (2020). Structuring decisions under deep uncertainty. *Topoi*, 39(2):257–269.
- Hugoson, A., Koch, G., Göthberg, C., Helkimo, A. N., Lundin, S.-A., Norderyd, O., Sjödin, B., and Sondell, K. (2005). Oral health of individuals aged 3–80 years in jönköping, sweden during 30 years (1973–2003). ii. review of clinical and radiographic findings. *Swedish Dental Journal*, 29(4):139–155.
- Huntington, D. and Lyrantzis, C. (1998). Improvements to and limitations of latin hypercube sampling. *Probabilistic engineering mechanics*, 13(4):245–253.
- Ibrahim, H. (2017). Car parking problem in urban areas, causes and solutions. In *1st International Conference on Towards a Better Quality of Life*.
- Jaxa-Rozen, M. and Kwakkel, J. (2018). Tree-based ensemble methods for sensitivity analysis of environmental models: A performance comparison with sobol and morris techniques. *Environmental Modelling & Software*, 107:245–266.
- Kasprzyk, J. R., Nataraj, S., Reed, P. M., and Lempert, R. J. (2013). Many objective robust decision making for complex environmental systems undergoing change. *Environmental Modelling & Software*, 42:55–71.
- Kwak, C. and Clayton-Matthews, A. (2002). Multinomial logistic regression. *Nursing research*, 51(6):404–410.
- Kwakkel, J., Walker, W., and Marchau, V. (2012). Assessing the efficacy of adaptive airport strategic planning: Results from computational experiments. *Environment and Planning B: Planning and Design*, 39(3):533–550.
- Kwakkel, J. H. (2019). A generalized many-objective optimization approach for scenario discovery. *Futures & Foresight Science*, 1(2):e8.
- Kwakkel, J. H. (2021). Open exploration. [https://emaworkbench.readthedocs.io/en/latest/indepth\\_tutorial/open-exploration.html](https://emaworkbench.readthedocs.io/en/latest/indepth_tutorial/open-exploration.html).
- Kwakkel, J. H. and Haasnoot, M. (2019). Supporting dmdu: A taxonomy of approaches and tools. In *Decision Making under Deep Uncertainty*, pages 355–374. Springer, Cham.
- Landry, M., Malouin, J.-L., and Oral, M. (1983). Model validation in operations research. *European Journal of Operational Research*, 14(3):207–220.
- Lempert, R. (2007). Can scenarios help policymakers be both bold and careful? *Blindside: How to Anticipate Forcing Events and Wild Cards in Global Politics*, Brookings Institution Press, Washington DC.
- Lempert, R., Nakicenovic, N., Sarewitz, D., and Schlesinger, M. (2004). Characterizing climate-change uncertainties for decision-makers. *Climatic Change*, 65(1-2):1–9.
- Lempert, R. J. (2002). A new decision sciences for complex systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7309–7313.

- Lempert, R. J. (2003). Shaping the next one hundred years: new methods for quantitative, long-term policy analysis.
- Lempert, R. J. (2019). Robust decision making (rdm). In *Decision making under deep uncertainty*, pages 23–51. Springer, Cham.
- Lempert, R. J., Bryant, B. P., and Bankes, S. C. (2008). Comparing algorithms for scenario discovery. *RAND, Santa Monica, CA*.
- Lempert, R. J., Popper, S. W., Groves, D. G., Kalra, N., Fischbach, J. R., Bankes, S. C., Bryant, B. P., Collins, M. T., Keller, K., Hackbarth, A., et al. (2013). Making good decisions without predictions: Robust decision making for planning under deep uncertainty. *RAND Corporation, Santa Monica, California*. Available: [rand.org/pubs/research\\_briefs/RB9701](http://rand.org/pubs/research_briefs/RB9701). (March 2016).
- Liu, Q. and Homma, T. (2009). A new computational method of a moment-independent uncertainty importance measure. *Reliability Engineering & System Safety*, 94(7):1205–1211.
- Lowe, S., Morrell, D., and Copley, G. (1982). Uncertainties in highway appraisal: the development of systematic sensitivity testing. In *Planning & Transport Res & Comp, Sum Ann Mtg, Proc*.
- Lyons, G. and Marsden, G. (2019). Opening out and closing down: the treatment of uncertainty in transport planning's forecasting paradigm. *Transportation*, pages 1–22.
- McMichael, A. J., Campbell-Lendrum, D. H., Corvalán, C. F., Ebi, K. L., Githeko, A., Scheraga, J. D., and Woodward, A. (2003). *Climate change and human health: risks and responses*. World Health Organization.
- McNally, M. G. (2007). *The four-step model*. Emerald Group Publishing Limited.
- McPhail, C., Maier, H., Kwakkel, J., Giuliani, M., Castelletti, A., and Westra, S. (2018). Robustness metrics: How are they calculated, when should they be used and why do they give different results? *Earth's Future*, 6(2):169–191.
- Ottomanelli, M. and Wong, C. K. (2011). Modelling uncertainty in traffic and transportation systems.
- Overtoom, I., Correia, G., Huang, Y., and Verbraeck, A. (2020). Assessing the impacts of shared autonomous vehicles on congestion and curb use: A traffic simulation study in the hague, netherlands. *International journal of transportation science and technology*, 9(3):195–206.
- Peng, J., Xiao, H., Wang, J., and Zhang, J. (2020). Impact of severe smog on travel demand of residents in tourist generating places: a case study of beijing. *Current Issues in Tourism*, 23(16):2009–2026.
- Petrik, O., Adnan, M., Basak, K., and Ben-Akiva, M. (2018). Uncertainty analysis of an activity-based microsimulation model for singapore. *Future Generation Computer Systems*.
- Planbureau, C. and voor de Leefomgeving, P. (2015). Cahier mobiliteit. <https://www.cpb.nl/sites/default/files/publicaties/bijlagen/cpb-pbl-boek-19-wlo-2015-mobiliteit.pdf>.
- Rasouli, S. and Timmermans, H. (2012). Uncertainty in travel demand forecasting models: literature review and research agenda. *Transportation letters*, 4(1):55–73.
- Ratrouf, N. T. and Rahman, S. M. (2009). A comparative analysis of currently used microscopic and macroscopic traffic simulation software. *The Arabian Journal for Science and Engineering*, 34(1B):121–133.

- Rijksoverheid (2020). *Rapportage Rijkswegennet*. <https://www.rijksoverheid.nl/documenten/rapporten/2020/10/26/rapportage-rijkswegennet-t2-2020>.
- Rijkswaterstaat (2021). Verkeers- en vervoersmodellen lms en nrm. <https://www.rijkswaterstaat.nl/wegen/wegbeheer/aanleg-wegen/nederlands-regionaal-model-nrm-en-landelijk-model-systeem-lms>.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.
- Schrijver, J., Meeuwissen, A., and Hilbers, H. (2003). Het voorspellen van betrouwbaarheid van reistijd met een vevoerprognosemodel: Smara, proceedings of the cvs.
- Sheahan, M. (2018). German court bans older diesel cars in parts of berlin. <https://www.reuters.com/article/us-germany-emissions-idUKKCN1MJ18N>.
- Sheikholeslami, R. and Razavi, S. (2017). Progressive latin hypercube sampling: An efficient approach for robust sampling-based analysis of environmental models. *Environmental modelling & software*, 93:109–126.
- Van der Heijden, K. (2011). *Scenarios: the art of strategic conversation*. John Wiley & Sons.
- van der Scheer, F. (2021). Compound risk: when sudden-onset disasters coincide with covid-19: Balancing the livelihood and the covid-19 trajectory of rural communities in developing countries.
- van Eck, J. R., Euwals, R., and Hilbers, H. (2020). Past “corona” in de bandbreedte van de wlo?
- Walker, W. E., Lempert, R. J., and Kwakkel, J. H. (2012). Deep uncertainty. *Delft University of Technology*, 1(2).
- Walker, W. E., Marchau, V. A., and Kwakkel, J. H. (2013). Uncertainty in the framework of policy analysis. In *Public Policy Analysis*, pages 215–261. Springer.
- Walley, P. (1991). *Statistical reasoning with imprecise probabilities*.
- Yang, B., Wu, Y.-g., and Ren, B. (2012). Application of multi-resolution modelling in emergency evacuation simulation. *International Journal of Simulation and Process Modelling*, 7(1-2):89–97.
- Yang, Q. and Morgan, D. (2006). Hybrid traffic simulation model. Technical report.
- Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.
- Zhao, Y. and Kockelman, K. M. (2002). The propagation of uncertainty through travel demand models: an exploratory analysis. *The Annals of regional science*, 36(1):145–163.



## POSSIBLE MEASURES

Scenario	Autonoom	Laag scenario	Hoog scenario (excl. Auto)	Hoog scenario <b>NB met en zonder OHT</b>
<b>Knop</b>				
Auto	Aanpassing conform bestaand OT model (check Ring W noordelijk deel, zo laten als in OT)	30 km/u maatregelen + 4 knips + autoverbinding Oosterhamriktracé (OHT) Veranderingen: Ring West ongelijkvloers	30 km/u maatregelen + 4 knips + autoverbinding Oosterhamriktracé (OHT)	30 km/u maatregelen + 10 knips + OHT + aansl. de Ruyterlaan op Hoendiep + volledige aansl. Iepenlaan
OV	Vastgestelde Dienstregeling bus 2021 (DB 2 oktober 2020) + Dienstregeling trein 2023 Veranderingen DB 2020 – DB 2021	gewijzigde lijnvoering minimaal, dienstregeling gemiddeld <b>Scenario 1</b> <ul style="list-style-type: none"> <li>• Q-liner 8x</li> <li>• Q-link 8x muv Q-link 15 6x</li> <li>• Tangent 6x muv lijn 8 2x</li> <li>• Dragend Stad / Streek 4x</li> <li>• Ontsluitend Stad / Streek 2x muv 178, 107/109 4x</li> </ul> Structuur: zie PPT (kaarten!)	gewijzigde lijnvoering basis, dienstregeling hoog <b>Scenario 2</b> <ul style="list-style-type: none"> <li>• Q-liner 10x</li> <li>• Q-link 12x muv Q-link 15 8x</li> <li>• Tangent 8x muv lijn 8 4x</li> <li>• Dragend Stad / Streek 6x</li> <li>• Ontsluitend Stad / Streek 4x</li> </ul> Structuur: zie PPT (kaarten!)	gewijzigde lijnvoering basis, dienstregeling hoog <b>Scenario 2</b> <ul style="list-style-type: none"> <li>• Q-liner 10x</li> <li>• Q-link 12x muv Q-link 15 8x</li> <li>• Tangent 8x muv lijn 8 4x</li> <li>• Dragend Stad / Streek 6x</li> <li>• Ontsluitend Stad / Streek 4x</li> </ul> Structuur: zie PPT (kaarten!)
Fiets	Fietsstrategie	eindbeeld fiets	eindbeeld fiets	eindbeeld fiets
Anders reizen	Geen verandering	HBO/WO 10% buiten spits MBO 5% buiten spits Dienstensector 10% buiten spits, 0,5 dag minder/week Overheidssector 10% buiten spits, 0,5 dag minder/week Eenmanszaak voor 25% conform dienstensector	HBO/WO 40% buiten spits MBO 10% buiten spits Dienstensector 20% buiten spits, 1,5 dag minder/week Overheidssector 20% buiten spits, 1,5 dag minder/week Eenmanszaak voor 25% conform dienstensector	HBO/WO 40% buiten spits MBO 10% buiten spits Dienstensector 20% buiten spits, 1,5 dag minder/week Overheidssector 20% buiten spits, 1,5 dag minder/week Eenmanszaak voor 25% conform dienstensector
Parkeren	Betaald parkeren oranje gebieden met aangepaste tarieven	hele stad betaald parkeren = oranje + geel	hele stad betaald parkeren incl. De Wijert + Vinkhuizen, terugdringen bewonersparkeren door aantal plaatsen, minder autobezit en vergunningensysteem	hele stad betaald parkeren incl. De Wijert + Vinkhuizen, terugdringen bewonersparkeren door aantal plaatsen, minder autobezit en vergunningensysteem
HUB/P&R	Traditioneel	Huurfietsen, 12 km/u, bereik tot 3 km (bestemmingskant)	Snel huurvervoer (E-bike, scooter, auto) 20 km/u, bereik tot 5 km (bestemmingskant), winkels bij P+R	Snel huurvervoer (E-bike, scooter, auto) 20 km/u, bereik tot 5 km (bestemmingskant), winkels bij P+R

Figure A.1: Possible policies



# B | FEATURE SCORING

A feature scoring is conducted for every policy. This is to show the effect of policies on the effect of uncertainties of policies. Added to this is the 95% confidence interval for car\_total

## B.1 30 KM POLICY

Results are found in figure B.1 and B.2.

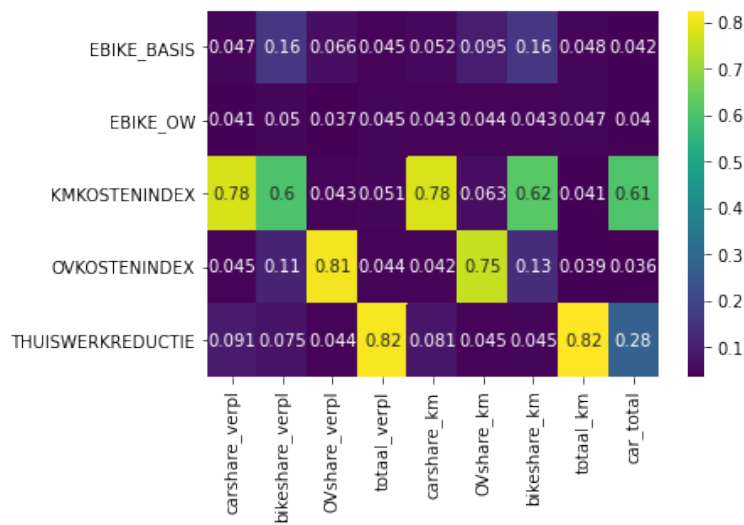


Figure B.1: Feature scoring with the 30 km policy

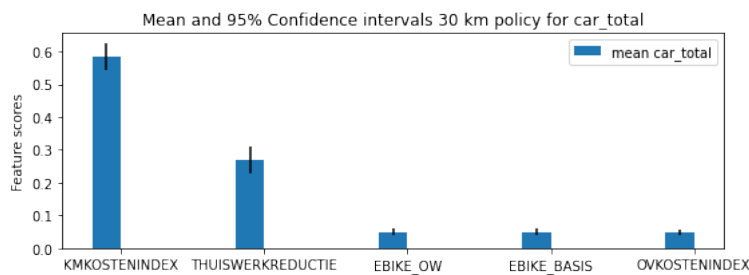


Figure B.2: 95% confidence interval of the feature scoring of car\_total for the 30 km policy

## B.2 BIKE POLICY

Results are found in figure B.3 and B.4.



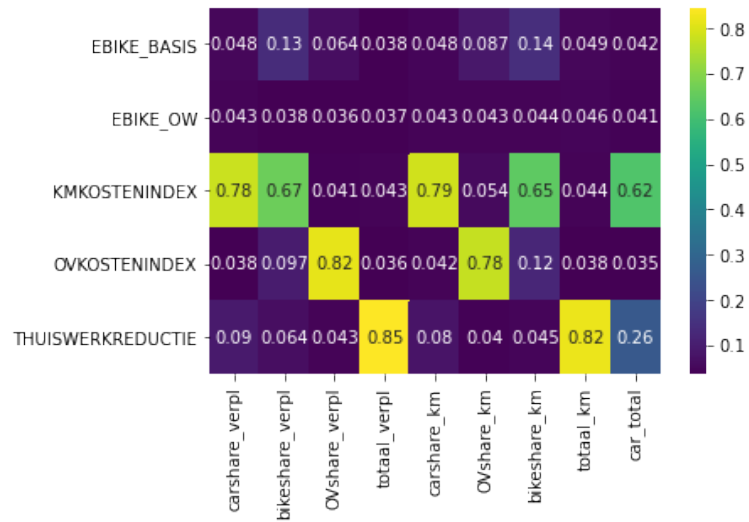


Figure B.3: Feature scoring with the bike policy

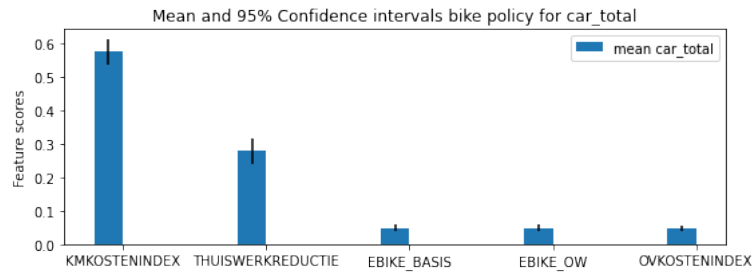


Figure B.4: 95% confidence interval of the feature scoring of car\_total for the bike policy

### B.3 PARKING POLICY

Results are found in figure B.5 and B.6

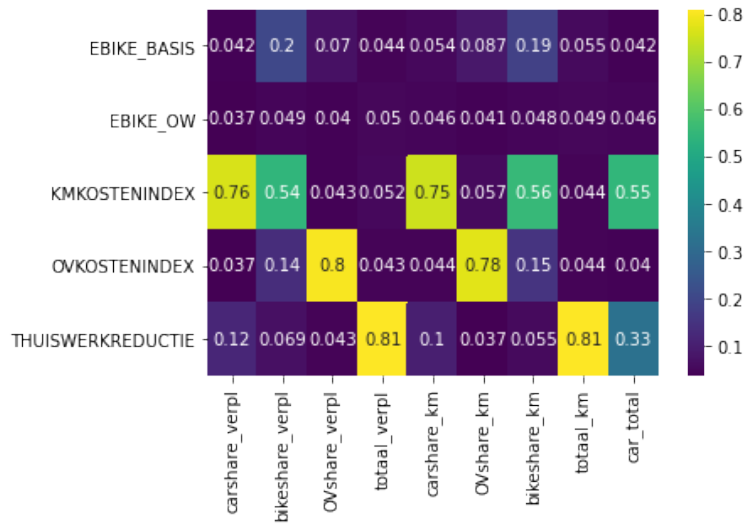


Figure B.5: Feature scoring with the parking policy

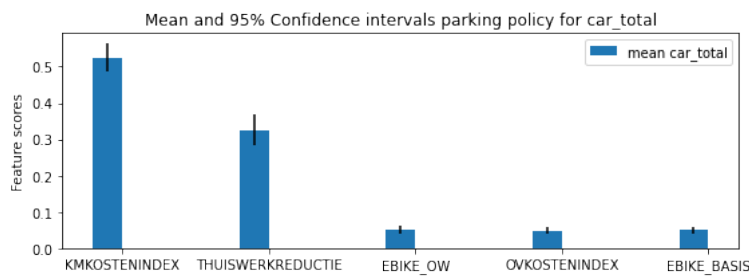


Figure B.6: 95% confidence interval of the feature scoring of car\_total for the parking policy



## C.1 PEELING AND COVERING PROCESS

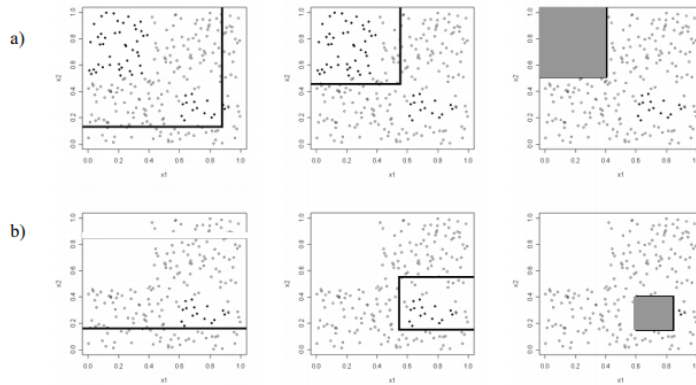


Figure C.1: Conceptual illustration of PRIM's: a) peeling and b) covering process. (Lempert et al., 2008)

## COLOPHON

This document was typeset using  $\text{\LaTeX}$ . The document layout was generated using the `arsclassica` package by Lorenzo Pantieri, which is an adaption of the original `classithesis` package from André Miede.

