Assignment 4: Exploratory Model Analysis

Matthijs Brouns 1505157 18th of December, 2014

Hélène van Heijningen 4076850

*Analyze the model you have built carefully or choose to use the 2-area epidemic model supplied. What are the main uncertain factors (7 to 15 factors) in this model? What are justifiable ranges/values for these factors? Use the sensitivity analysis tools in Vensim to explore the behavior of your model over these ranges. Analyze the results and discuss the policy implications of your findings.*

*In your write up of the assignment, clarify which uncertain factors you have chosen and why, what ranges you have used and why, motivate your sampling technique, and your analysis of the results. The write up should be such that your analysis can be replicated by me, in theory.*

*Tips:*

* *You can export the data from the sensitivity analysis to tab separated or comma separated files for further analysis. Read the Vensim manual for details*
* *Many of the discussed techniques (PRIM, CART) are available in open source implementations. You could even use the code of the exploratory modeling workbench if desired.*
* *Ranges should be plausible and grounded in information.*
* *Reference the theory material.*

# Approach

In this Section we will explain our approach to exploratory model analysis of our version of the Kirkwood water crisis model.

We start off by identifying the most uncertain yet important parameters in our model. We identified these parameters by looking at the structure of the model, using the causal diagram shown in figure 1, as well as by using the results from the sensitivity analysis that has been performed earlier. After going through the different parameters and structures that are uncertain in the model, we have chosen to focus on 10 factors. Why we chose to take into account certain factors and why we decided to keep others out will be explained in the next Section.

After the factors were chosen, their ranges had to be determined. These had to be realistic, meaning that the values within those ranges might occur in the real system even if they happen in uncommon scenarios. Doing this for all the different factors was difficult as it involves a lot of guessing, but we will explain more on that in the Section on factors and their ranges.

Next we determined what output variable we wanted to do the multi-variate analysis for. This is discussed in the Section about classifying model behaviour.

The multi-variate analysis was then run and the data was loaded into R to do further analysis. We wanted to classify the different types of behaviour produced by the multi-variate analysis to then use this to find out what values the factors had to adopt so that the model would give the behaviour we were looking for; Kirkwood where no crisis happens.

Based on the results found from the multi-variate analysis (factor x, y and z need to take on values equal to d, e and f for example), policies can be created. Then the robustness of these policies is tested by implementing the policies into the model to see under what circumstances they hold.

This is just a short summary of how we approached this assignment. A more elaborate explanation will follow in the different sections.

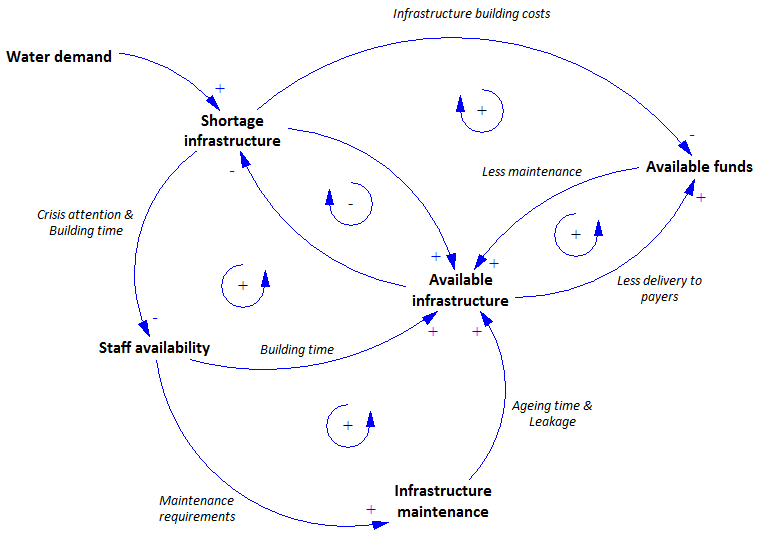


Figure 1: Causal model of the Kirkwood water infratructure system

# Factors and their ranges

When determining the factors that were uncertain and needed further exploration with regards to the possible effects of these factors on the model behaviour, we identified 10 factors. Of course there are many more factors in our model and we will now explain why we chose to take these into account and keep other factors out of this exploratory model analysis.

## Staff availability

From the sensitivity analysis earlier performed we knew that factors influencing the staff availability had a relatively large effect on the model behaviour, especially on whether a crisis would occur or not. The factors that influence the staff availability are the number of staff hired, hiring time period and the staff service time. From these factors we chose to take into account the number of staff hired per time period (so the hiring time period stays the same, but we change the number of people hired) and the staff service time because those are the main causes of changes in staff availability. In terms of policies we also thought these two factors would be interesting as a policy could consist of both hiring more people, or educating more people so there are more suitable workers, and of making a staff member’s service time longer by providing a more attractive working environment and agreements.

### Staff hired / time period

Range: 0 – 25 staff members hired per 2 years

Distribution: Vector (0;25;1)

The number of staff hired per time period (2 years) has a base value in the model of 5 members per 2 years. We have set this range in the multi-variate analysis from 0 to 25 staff members hired per 2 years. 25 per 2 years may be too much for the current system to be able to do in the future, but we wanted to be sure that we would grasp the effect of this very important factor in its entirety, even if this meant that the model might only show no crisis when the staff hired per 2 years had to increase to 25 members. This has been modeled as a vector incremented by 1, since you cannot have 3.4 staff members for example.

### Staff service time

Range: 0 – 25 years

Distribution: Uniform (0;25)

The staff service time has a base model value of 10 years (it takes 10 years on average for a staff member to leave). We have set the range for this factor from 0 to 25 years as this seemed a realistic value to where the staff service time could increase, by for example implementing a policy that provides better work agreements and a better environment for the workers so they will stay longer. The distribution that has been used is a uniform distribution between 0 and 25 so that all possible time periods are taken into account.

## Distribution of staff attention

The other part of the staffing system involved dividing the attention of available staff over the different activities such as maintenance, planning, constructing and refurbishing activities. Factors that influence this system are the average production rate of a staff member, the distribution of activities between primary and secondary activities, the distribution of activities between planning and construction tasks (secondary activities) and finally the number of staff required for one tanker. We have decided to take the last three factors into account and leave the first one, the average production rate per staff member out of this analysis. This is because the average production rate per staff member is not an uncertain value and could only theoretically be altered by making the work of the staff members more simple so they can do more in the same time period. However, we believe this is not a policy that is easily comprehensible or implementable and above all not realistic as there is no money available in Kirkwood to set up a more efficient working environment. The other factors however are very interesting as they control how much staff attention is given to different tasks. By altering this distribution we are curious to see if this could have a positive impact on the system.

### Attention rate for primary / secondary staff activities

Range: Proportionally; Primary activities first; Secondary activities first

Distribution: Vector (0;1;0.5)

The attention rate for primary / secondary staff activities is currently being distributed proportionally to the amount of maintenance production required and secondary staff production required. This distribution could however be changed by altering the model structure. In the multi-variate analysis we include this factor by including a variable in the model which changes the model structure when it takes on one of the following three values: 0, 0.5 or 1. We distinguish between three possible structures:

* Divide staff by ratio between primary and secondary activities
* Do all maintenance tasks first before giving attention to secondary activities
* Do all secondary tasks first before giving attention to primary activities

### Attention rate between planning and construction tasks

Range: Between 1% planning, 99% refurbishment and 99% planning, 1% refurbishment

Distribution: Uniform (0.01;0,99)

This factor has a base model value that distributes the planning and refurbishment tasks as follows: 25% of the attention of secondary tasks is given to planning and 75% of the attention is given to refurbishment activities. To see what effect this uncertain factor could have on the model behaviour it has been taken into account with a uniform distribution of values between 1% and 99% to make sure that at least some attention is given to all activities at all times.

### Staff required for tankers

Range: 0 – 1 staff members per tanker

Distribution: Vector (0;1;1)

In our model we made the assumption that part of the staff that works on infrastructure would have to work on supplying water to the households when tankers were in play in case of a crisis. This means that when there is a crisis and tankers are needed, a large part of the entire group of available staff members is taken away from infrastructure jobs and towards the tankers. Since this is an assumption on our part, this is an uncertain factor which could influence how crisis would evolve during time. Therefore it was very important to take this factor into account. The range of this factor is set to either 0 staff members being needed to tankers or 1 staff member needed, equal to the current model setup.

## Funding system

The funding system is the next system in which several uncertain factors can be found. The main factors influencing this system are the annual financial bail-out, the percentage of income used for maintenance, the price per unit of infrastructure, the maintenance costs per unit infrastructure and the percentage of delivery costs recovered from households and other urban users. The annual financial bail-out had already been established as a very uncertain factor as the government was already contemplating removing this subsidy in the future. Because of this, this factor is taken into account to get a grasp of the real importance of this factor in the system. Furthermore, the percentage of income used for maintenance is taken into account to see if a different budget allocation of the Kirkwood municipality might influence the system for the better. The maintenance and infrastructure unit costs are not taken into account as these factors are not uncertain. Lastly, the percentage of delivery costs recovered has not been taken into account, although this is an interesting factor which could have a large influence on the system. We have however chosen not to take this factor into account since designing a policy to influence people to register etcetera so the bills do get paid is extremely difficult, time consuming and expensive without the certainty that it will actually work. For this reason we prefer to focus on other policies that are easier to implement and have a higher chance of success.

### Annual financial bail-out

Range: 0 – 200 kZAR per year

Distribution: Uniform (0;200)

The model base value for this factor is 150 kZAR per year. Given the introduction on the Kirkwood situation, we were told that the government might remove this subsidy in the future. Knowing this, we found it most realistic to set the range of this factor from 0 kZAR per year, completely removing the subsidy, to 200 kZAR per year which would mean a slight increase. This increase is much less likely to happen in real life compared to the value of 0, however we wanted to take into account the chance that increasing this subsidy might solve a lot of Kirkwood’s problems. A uniform distribution is chosen so all sizes of the subsidy are explored.

### Percentage of income used for maintenance

Range: 20% to 60%

Distribution: Uniform (0.2;0.6)

Currently the value in the model is set at 40%. We have decided that this value can be altered between 20% and 60%, so the amount of money used for maintenance remains within a realistic range since using more than 60% of your income for infrastructure projects will of course create problems in other areas of the Kirkwood budget. A uniform distribution is again used so all values in between 20% and 60% can be investigated.

## Infrastructure system

The last system which is based on uncertain factors is the infrastructure system. The infrastructure system includes quite some uncertain factors, such as the leakage percentage, the infrastructure aging time, the maximum percentage of infrastructure pushing, the target percentage of infrastructure maintenance, the planning time, the building time and ???. We have identified three of these factors that we want to take into account in this analysis. These are the infrastructure aging time, the maximum infrastructure pushing and the target percentage infrastructure maintenance. We chose to take into account the infrastructure aging time since the sensitivity analysis showed that this part of the model has quite a large influence on the system. This factor in itself is not easy to influence in the real system, unless a new type of infrastructure was used with a longer / shorter time span. Even though this might not be the most realistic policy, we wanted to take into account this factor since we expect it to have quite a large influence on the system which we can now investigate further. The maximum infrastructure pushing is the percentage showing how much the infrastructure can be pushed. Infrastructure pushing causes infrastructure to age faster, but it also creates a buffer in the system which can prevent crises. This is thus an interesting factor as it affects a large part of the system, which is why we chose to investigate this factor in this analysis. The last factor we chose to include is the target percentage infrastructure maintenance. This is the percentage of infrastructure that the system wants to maintain if it has enough staff to do this. By maintaining more infrastructure less new infrastructure needs to be built since the life span of infrastructure goes up, the opposite happens when less is maintained. This is a very important factor in the system and altering this value might have a great influence on whether crises happen or not.

The leakage, planning time and building time are not taken into account as these factors are not very uncertain. Neither are the values unnecessarily high or unrealistic, so we chose to keep these factors out of the analysis.

### Infrastructure aging time

Range: 20 – 40 years

Distribution: Uniform (20;40)

The base model value of this factor is 30 years. We have decided to give this factor a range between 20 and 40 years, so the life span will remain a realistic value. A uniform distribution is chosen so the entire area between 20 and 40 years is explored.

### Max infrastructure pushing

Range: 0% - 50%

Distribution: Uniform (0;0.5)

The base model value of this factor is set at 30%. This means the infrastructure is used for 130% of its capacity, thereby decreasing the lifespan of the infrastructure. As stated earlier we are interested to see how changing this ‘buffer’ can influence the system. To do this a realistic range has been chosen from 0% (no buffer) to 50%, where the infrastructure will be pushed immensely. We chose to cap this value at 50% because pushing the infrastructure even more would become unrealistic and the life span of infrastructure would drastically decrease, crashing the system.

### Target percentage infrastructure maintenance

Range: 5% - 25%

Distribution: Uniform (0.05;0.25)

The base value of this factor in the model is equal to 8%. We have chosen to set the range of this factor in this analysis between 5% and 25%, thereby testing an even lower amount of infrastructure maintenance and a much higher target infrastructure maintenance. We believe these values represent a realistic range of the target percentage of infrastructure maintenance, ranging from only maintaining a small amount of infrastructure to maintaining 25% of the infrastructure. Maintaining more than 25% of the infrastructure would seem very unrealistic as this would require too much staff members and above all, would not be necessary.

# Model behavior classification

In order to increase the interpretability of the multi-variate analysis outcomes, the resulting behaviours need to be classified. We have explored several classification methods which will be described in this section.

Choosing the classification method and variable is based on three factors. First, the number of classes needs to be limited. Having a high number of classes or even multiple dimensions of classes greatly increases the number of runs needed to draw statistically valid conclusions on the behaviour since otherwise the model runs will be distributed too sparse over the output space. This effect is often referred to as the curse of dimensionality. The second factor is that the classes should offer a good insight into the actual behaviour of the model. The last factor is regarding the interpretability of the classified outcome. The results should be easily interpretable by a stakeholder.

### Classify behavior on type of crisis which occurs

The first classification method we explored was the classification of the model behaviour on the type of crisis which occurs. There are a number of different possibilities for crisis types:

* Infrastructure crisis
* Staffing Crisis
* Financial Crisis
* No crisis

A main advantage of this classification approach is that it offers a low number of classes and that the different classes can be easily understood. However, this classification method is deemed unsuitable for this model because of two reasons.

First of all, the classification offers no insight into the severity of the crisis. It also gives no information on whether the crisis eventually gets solved.

Furthermore, exploration of the model showed that the different crises often follow up on eachother rather quickly. A financial or staffing crisis will almost always cause an infrastructure crisis. An infrastructure crisis will also almost always eventually cause a financial crisis.

### Classify behavior on timeslot where crisis first occurs

The second classification method is based on the timeslot in which a crisis occurs for the first time. The time-slots can for example be:

* 0-10 years
* 10-20 years
* 20-30 years
* 30-40 years
* 40-50 years

Once again, the interpretability of these classes is rather high. There is also a limited number of classes and only in one dimension.

A disadvantage of choosing this classification method is that there is no indication on the severity of the crisis and it is not determined whether the crisis eventually gets solved.

Therefore, this classification method is not chosen for this experiment.

### Classify behavior on number of years of crisis

The third classification method which was explored is based on the number of years in which there was a crisis. Once again this leads to a low number of classes. A major disadvantage is that there is no indication of the severity of the crisis and it is not known when the crisis occurs.

### Classify behavior on number of years of crisis + timeslot where crisis occurs

When we combine the previous two classifiers into a two-dimensional class consisting of both the timeslot of the crisis and the number of years in which there was a crisis some of the problems of the previous two classifiers are solved.

The main problem with using this classifier is that it is a two-dimensional output space. This massively increases the number of output possibilities and therefore the amount of runs needed to draw statistically valid conclusions. Furthermore, two-dimensional output spaces are much harder to interpret for humans since it becomes difficult to say whether a single class is better than another class.

### Classify behavior on infrastructure shortage \* period (area under the curve)

Finally we settled on a classifier which uses the area under the curve of the infrastructure shortage graph. This area clearly indicates the severity of the crisis since it consists of both the infrastructure shortage above the amount that is being pushed, and the period in which there was a shortage. Therefore, this classifier gives very accurate and high resolution results.

A disadvantage of this method is that it is slightly more difficult to interpret than the other classifiers since ML\*Year is not an easily understood unit. To mitigate this, the value is divided by the average household water need to result in a Household\*year water shortage figure which is already much better interpretable. After the multivariate analysis had been performed only the values at year 50 (end of model run) of this figure were saved, since this figure on its own already indicates whether there has been a crisis or not. Since we wanted to explore what values the factors had to adopt in order to create model behaviour in which there was no crisis at all, only saving the value of our output variable at year 50 gave us enough insight to do our analysis.

Something which needs to be kept in mind is that this choice of classifier means that a low number of households with a long period of shortage is classified the same as many households with short shortages. This may or may not be desirable for all purposes.

Since we are mostly interested in the runs which produce a low area under the curve, the bin sizes are not equal. The used bins are:

* 0 household\*year
* 0-1,500 household\*year
* 1,500-3,000 household\*year
* 3,000-6,000 household\*year
* 6,000-9,000 household\*year
* 9,000-12,000 household\*year
* 12,000-20,000 household\*year
* 20,000-30,000 household\*year
* 30,000-40,000 household\*year
* 40,000-50,000 household\*year
* 50,000-100,000 household\*year
* 100,000+ household\*year

For this classification method we chose to run 10,000 runs. This results in about 1,000 model runs per class if all classes are distributed evenly over the solution space. This should provide enough data to infer results from while still being feasible to run in the limited timeframe of this assignment.

Because of the limited number of runs, a Latin Hypercube was used as a sampling method to ensure a equal distribution over the input space.

# Results

This section will describe the results of the Exploratory Model Analysis using the setup as described earlier in this document.

The distribution of model runs over the earlier introduced classes can be found in figure 2. This shows that there is a high number of runs in which the model does not show a crisis situation (no infrastructure shortage). These runs are grouped in the bin which is called ‘0’, meaning all these runs had a value of 0 for the output variable at the end of the run. This result was very interesting to us as this meant that the system could indeed function properly. All we needed to find out now is what values the 10 factors had to take on to give these outcomes.

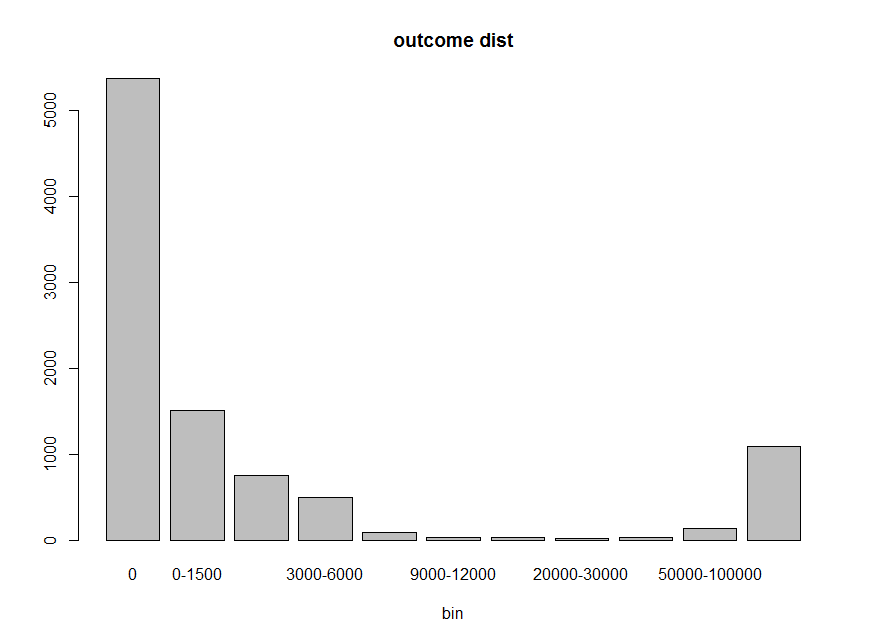


Figure 2: Histogram of the number of model runs distributed over several bins

Analyzing the effects of the change in input parameters on the model behaviour was done by using a decision tree. This decision tree was inferred using an C4.5 Decision Tree inference algorithm supplied by the RPart R Package. In order to avoid overfitting on the data, the tree was pruned optimizing on a lowest cross-validation error. The results of this decision tree are shown in figure 3.

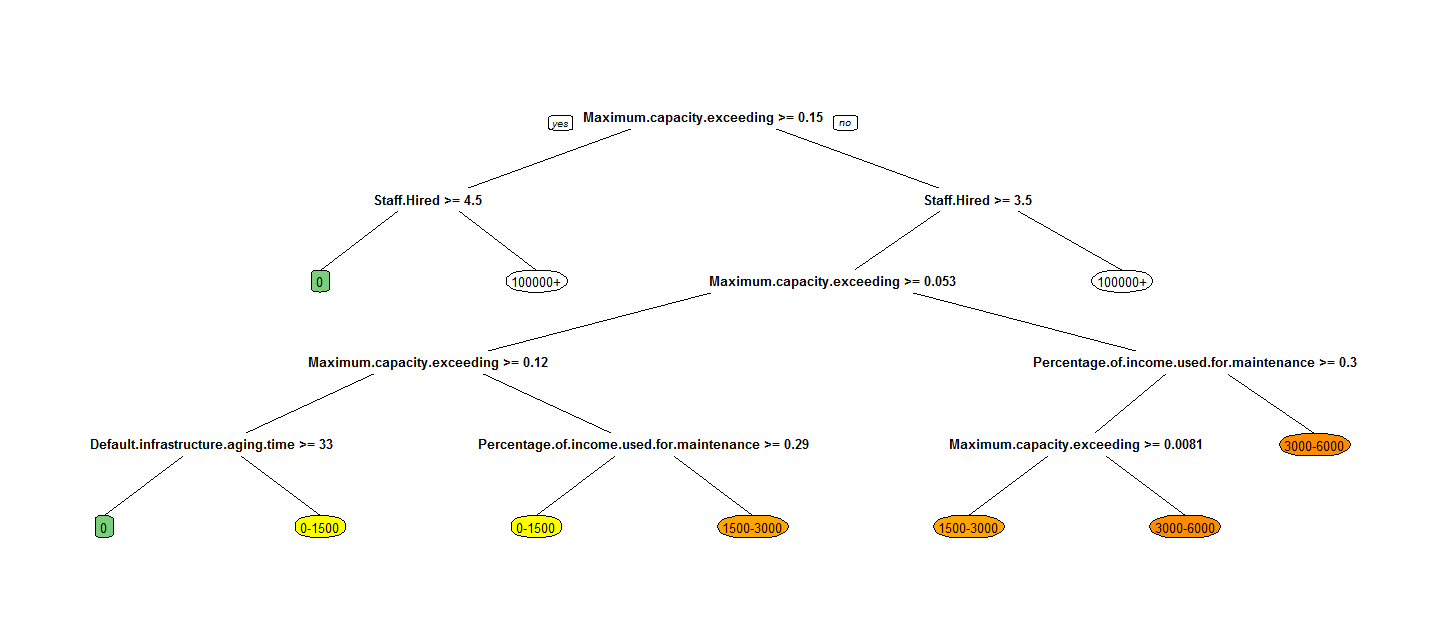


Figure 3: Decision tree resulting from the multi-variate analysis of the 10 chosen factors on the area beaneath the infrastructure shortage graph (number of households\*years of shortage)

# Policy

Before going into what policies we can infer from the decision tree shown in figure 3, we will shortly explain what can be seen in this decision tree. Since we want to explore what values the input variables need to take on so that the model does not experience a crisis, we need to climb down the decision tree towards the bin in which the output variable is equal to 0 at the end of the run. In the decision tree this is indicated by the green box with the ‘0’ inside of it. So, the decision tree shows at the top that when the value of the factor ‘Maximum capacity exceeding’ (percentage infrastructure pushing) is higher or equal to 0.15 and the number of staff hired per 2 years is equal or higher than 4.5 (so 5 people per 2 years) no infrastructure will occur in the system.

Then from the other side of the decision tree, when the maximum capacity exceeding is lower than 0.15 and the staff hired is higher than or equal to 4 people per 2 years and the maximum capacity exceeding is higher than 0.12 and finally when the default infrastructure aging time is equal or higher than 33 years the model also experiences no crisis situation.

From the decision tree in figure 3 we can then infer two policies in order to reach a system in which there are no crises occurring.

The first policy is to increase Staff hiring to 5 staff members per two years. Furthermore, a minimum of 15% infrastructure pushing should be kept available to act as a buffer.

In the second policy the staff hiring should be increased to 4 staff members per two years while keeping at least 12% and at most 15% infrastructure pushing available. Furthermore, infrastructure which has a slightly higher default lifespan of 33 years should be used instead of the regular 30 years.

In order to test the robustness of these policies a sensitivity analysis was performed using the parameters described in the first section. For the results of these sensitivity runs the same steps were performed regarding classification and decision tree building. The results of this sensitivity analysis for policy 1 is shown in figure 4. This decision tree shows that for this policy to work, the default lifespan of infrastructure should remain higher than 26 years. Furthermore, more than 32% of the municipalities’ income should be made available for maintenance and the staff service time should be greater or equal than 10 years.

Testing the policy robustness for policy 2 shows that in order for this policy to be effective, at least 30% of the municipalities’ income should be available for maintenance. Furthermore, staff service time should be greater than or equal to 12 years and the staff attention should be distributed pro rato between construction and maintenance tasks. The resulting decision tree from testing policy 2 is shown in figure 5.

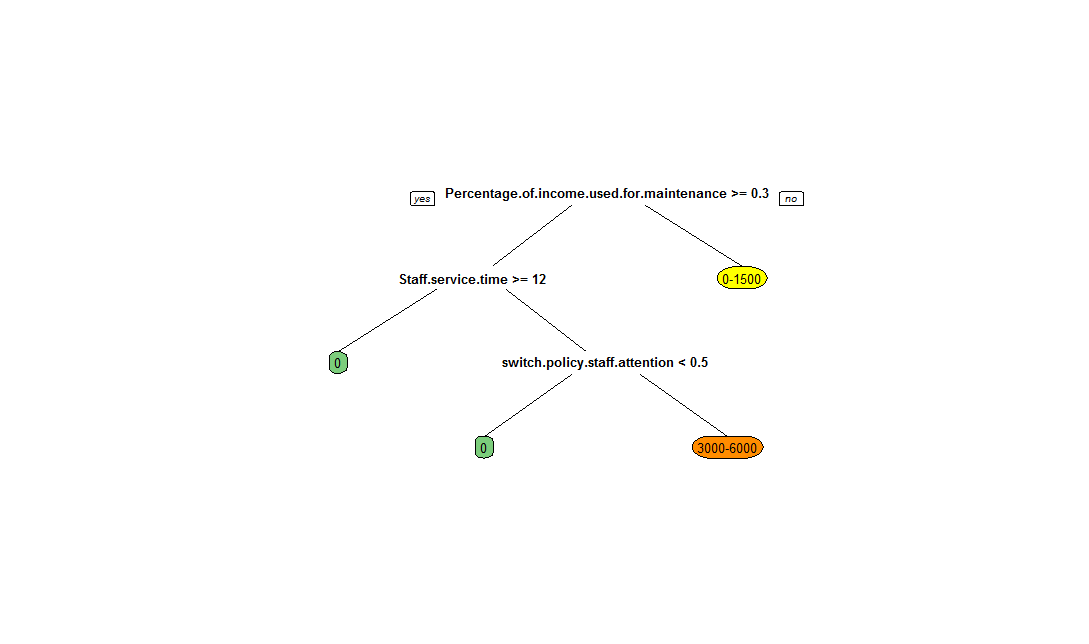
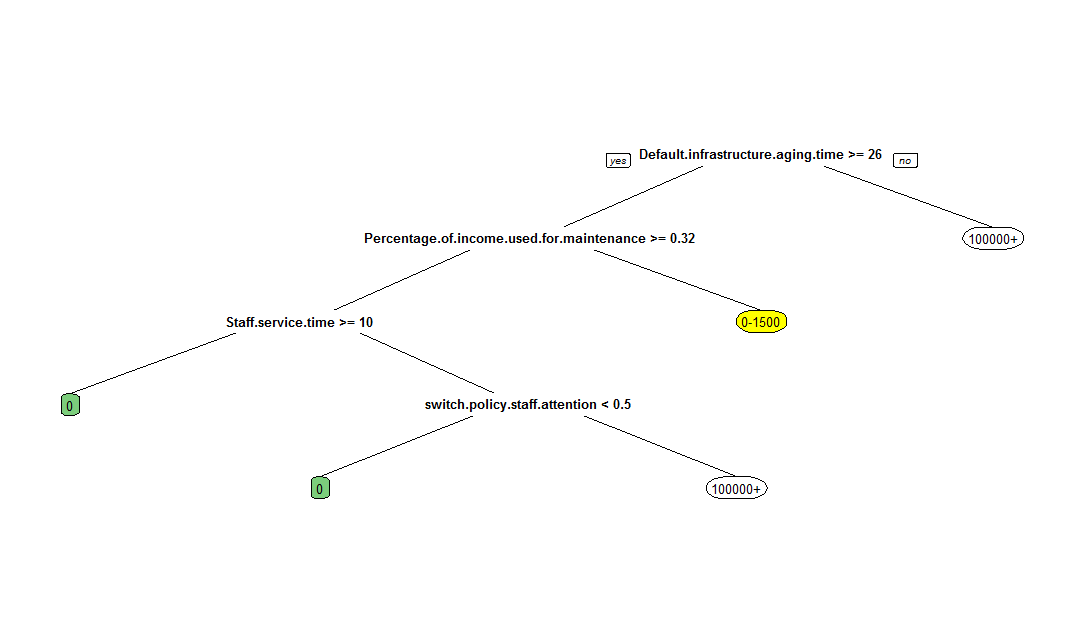


Figure 4: Decision tree resulting from a multi-variate analysis of the 8 remaining factors while the first policy has been implemented. The result is again the area under the infrastructure shortage curve

Figure 5: Decision tree resulting from a multi-variate analysis of the 7 remaining factors while the second policy has been implemented. The result is again the area under the infrastructure shortage curve

# Conclusion

The exploratory model analysis gave very positive results. It turned out there are indeed policies possible which will solve the systems issues so there are no more infrastructure shortages.

The first policy that has been identified includes increasing the number of staff hired from 2 to 5 members per 2 years and a infrastructure buffer should be kept available of minimally 15% infrastructure pushing. This policy is relatively easy to implement. In the current situation the infrastructure is already being pushed by 30% so keeping a minimum of 15% should not be a problem. The policy comes then down to increasing the hiring of staff so there is enough staff available to perform all infrastructure related tasks. The robustness of this policy is quite good since the decision tree in figure 4 shows all the values the factors need to adopt for this policy to hold which are equal to the current situation. Overall this policy is not too difficult to implement and the effects are very promising.

The second policy involves more changes to the current model values. The staff hiring should be increased to 4 staff members per two years while keeping at least 12% and at most 15% infrastructure pushing available. Furthermore, infrastructure which has a slightly higher default lifespan of 33 years should be used instead of the regular 30 years. Since the infrastructure pushing is currently already at 30%, that part of the policy shouldn’t be a problem. This leaves a doubling of the staff hiring variable and a slight increase in the lifespan of the infrastructure.

The doubling of the staff is just like in the first policy a feasible policy. However, the increase in the lifespan of the infrastructure would require infrastructure with a better quality. We are unsure whether this is feasible or not, but it is an interesting result to keep in mind.

Furthermore, the robustness of the second policy is more difficult as it only holds when the staff service time is increased from 10 to 12 years. A policy to achieve this could be about improving the working environment so people would want to work for a longer period of time or by improving the working agreements so staying is a better option for their future compared to leaving.

Overall we believe both policies are feasible, but we prefer the first policy because it is simpler to implement and it is robust under the given circumstances.

This analysis has been a great way to gain insight into the workings of our model and the system. We did not expect that we could achieve a model outcome without a crisis happening with only a relatively small change in the input variables.