Flood risk of the Helsinki road network

A network analysis determining travel time increases due to flooding in Helsinki

Name: Ewout ter Hoeven
Student number: 4493346
Date: 25-08-2021

Supervisors/advisors: Nazli Aydin & Ylenia Casali

Repository: https://github.com/EwoutH/Helsinki-flood-BEP-project

Abstract

In this study the flood vulnerability of the road network in Helsinki was researched. The main goal was to identify the highest risk road segments and measure the impact on the travel time of road users in different rainfall scenarios. Both the essentiality based on traffic patterns and the probability of a rainfall scenario heavy enough to cause the flooding of a road segment are included in this risk.

To accomplish this, the most used roads were identified based on shortest-path algorithms and origin-destination traffic data, using a modified weighted variant of betweenness centrality. A stormwater runoff model was constructed based on digital elevation models to be able to detect which roads would flood for a certain amount of rainfall. For 6 different scenario's – 4 regular and 2 with extreme values – the impact on total average travel time was measured, as well as the impact of the most used road segments.

Using the main traffic dataset, the mean travel time increases with 45.8%, 79.5%, 81.1% and 119.1% for the once in 10-, 20-, 50-, and 100-year rainfall scenarios respectively. The road segments contributing most to this increase were part of the Länsiväylä motorway (Western Highway), with the (150, 153)-segment alone adding 1.672% to the average trip time when flooded in the once in 20-year scenario. Many segments of the Kehä I highway (Ring I) also contribute heavily to the increase in traffic times, with up to 1.019% per segment starting at the once in 10-year scenario.

In future research, significant improvements to the stormwater model could be made by implementing multi-level structures such as highway overpasses, and including the remaining watershed volume of fully filled upstream bluespots when filling downstream bluespots. Traffic data could be improved by increasing the resolution and adding additional scenario's. The road graph could be improved by implementing reduced speed based on traffic volumes and representing intersections in more detail.

Table of contents

1. Introduction

Cities are prone to flooding with high levels of hard surfaces (like concrete) and low ability to drain water (Du et al., 2015, p. 1465). Climate change will worsen this by increased sea and river levels and more precipitation (Ruosteenoja et al., 2016). For the southern-Finland region, including Helsinki, an increasingly larger portion of the floods will be caused by heavy rainfall events, instead of snowmelt events which are currently the dominant cause of flooding (Veijalainen et al., 2010). This will impact many aspects of urban life, one of which is the ability to travel. Road transport will increasingly be disrupted by those floods (Pregnolato et al., 2017, p. 77), which is causing social and economic impact even on areas not directly affected.

Not all roads will flood at the same amount of rainfall however and the impact of a flooded road will be different for each of them. The City of Helsinki can modify road infrastructure to be more resilient to floods by increasing water drainage capacity or placing barriers on strategic points, but like most infrastructure projects these can get costly quickly. Therefore, it is useful to estimate the flood risk of each road so informed decisions can be made if certain infrastructure needs improved resilience and if so which specific road segments.

While general flood risk maps are available for Helsinki (Environment.fi, 2020), maps providing road risk, combining flood chance and impact, are not yet developed. These flood risk assessments have been done for other cities and countries, such as Bosnia and Herzegovina (Džebo et al., 2019), Bangladesh (Baky et al., 2019) and the Chang-Zhu-Tan Urban Agglomeration, China (Chen et al, 2019), but not yet Helsinki. Furthermore, the impact of each flooded road on the road network is not a static factor, but dependent on traffic demand patterns and the availability of surrounding roads, since there could be significant bottlenecking if essential roads get blocked. Different rainfall scenario's, with varying likelihood, will cause different road segment combinations to flood. High-level road network analysis has been done (in the public domain) for Helsinki, but only for specific cases (Juga et al., 2014) and not yet combined with flood impact.

Using spatial data a road graph will be distilled and processed for Helsinki and with water flow tracing a flood risk map will be constructed, based on a processed digital elevation model. Both models are linked based on the shared coordinate system to get a flood risk assigned. Shortest-path algorithms like Dijkstra's will be used to calculate the travel time between different origins and destinations, which will be weighted with origin-destination traffic data. These shortest paths and traffic amounts will be processed to calculate both traffic time delays, as well as the delay some of the most used road segments cause when flooded in one or more scenarios.

The aim of this study is to estimate the travel time impact of roads flooding in Helsinki. These models can be used by the City of Helsinki to improve the resilience of road infrastructure effectively and strategically.

2. Research

The main goal of this research is to map the weaknesses in the road system of Helsinki in different flood scenarios. To accomplish this, a research question is defined and split into specific sub-questions. The models and data used are briefly discussed in the conceptualisation, and methods are discussed briefly, before detailing the latter in the approach.

Research questions

The main research question is formulated as:

Which roads in Helsinki will disrupt traffic to the largest extent in different flood scenarios?

The research question can be further bisected into the following five sub-questions.

a) *Which roads in Helsinki, if unavailable due to flooding, lead to the largest decrease in regular traffic flow?*
   In regular use, which roads are essential in supporting daily traffic in Helsinki, and thus will have the largest **impact** when flooded either fully or partially. This impact can be either direct due to high water levels or indirect by congestion or preventive closure. These roads will be selected of which a graph of the most essential roads will be constructed.
b) *Of those essential roads in Helsinki, which are most prone to flooding?*
   Based on basic geometric and spatial data, like elevation profiles and surface permeability, which roads have the largest **probability** to be flooded? A statistical distribution of flood chance for different scenario's will be constructed and added the road graph.
c) *Which roads in Helsinki have the highest risk to disrupt traffic flow in different flooding scenarios?*
   Combining the regular traffic flow of each road, the potential reduction in maximum speed and the reduction in capacity based on the flood chance, which roads have the largest **risk** from floods? The road graph will be completed with this risk estimation, used for validation.
d) *By how much will the average travel time increase in different flood scenarios?*
   On a system level, by how much will the average travel time increase, based on simulations of the road graph? A discrete-event simulation model will be constructed, and the average travel time increase will be measured for each scenario.
e) *Which roads contribute most to this average travel time increase?*
   On a micro level, which roads are contributing the most to this average travel time increase? For each road, its impact will be measured, by randomly restricting or closing roads and measuring its impact.

Risk is defined here as the combination of probability and impact (Deloitte & Touche LLP et al, 2012). As can be seen in the first three sub-questions, the potential impact on traffic for each road will be measured, then the probability of flooding, based on the likelihood of the rainfall scenarios, and in the third sub-question those will be combined to a risk metric.

Sub-question d) and e) continue to measure the average travel time increase on macro level (per rainfall scenario and traffic dataset) and on micro level (per individual road segment).

Methods

To answer these questions, four main pieces of data are needed: Rainfall values per scenario, a digital elevation model, a graph describing the road network and traffic data.

Figure 2-1: Input-output diagram providing an overview of the primary data and analysis

The rainfall values and digital elevation model will be used to create flood risk maps, for which a modified version of the stormwater model as described by Balstrøm & Crawford (2018) is used. This model traces rainfall based on the elevation model to estimate where bluespots (which are depressions in the landscape where water can't flow out of) will emerge.

The model will be modified to be able to estimate not only how deep these bluespots are at different locations, but also how deep the water level is at each point for different rainfall levels.

Knowing the water depth for each location in different rainfall scenarios, these flood maps can be matched to the road graph based on a common coordinate system. For each road segment the average and maximum water depth is now known for each rainfall scenario.

Based on these water depth values; road segments can be removed from the road graph. The road graph contains travel time values for each segment as weights and using shortest-path algorithms like Dijkstra's the travel time between each combination of two locations can be found. Since not each combination of locations has the same amount of road traffic, traffic datasets containing origin-destination volumes will be used as weights to get a weighted average travel time for the whole road network.

To estimate the importance of each road segment, and thus the impact when a road segment is flooded, the concept of betweenness centrality is used to calculate the number of shortest paths that use each road segment. Since not each route sustains the same amount of traffic, the betweenness centrality also needs to be weighted with the origin-destination volumes from the traffic dataset. This new metric will be called OD-weighted betweenness centrality.

This OD-weighted betweenness centrality will be used to answer the last sub-question, the individual impact of each road segment. Road segments with the highest values will be removed one by one from the road graph and the increase in average travel time will be measured each time.

The exact steps performed in applying these methods are elaborated in the next chapter, the *Approach*, and its accompanying appendixes.

3. Approach

In this chapter the approach will be discussed, including the casus, input data used, model reductions, analysis methods and biases. This section is supported by Appendixes A, B, C and D, which provide more in-depth discussions.

Data assembly

The data for the models is selected to represent the real world as close as possible without making it computationally of otherwise unfeasible. Ockham's razor is applied when possible, to reduce unnecessary detail.

Road graph

The road graph is assembled from OpenStreetMap data. It was downloaded on 28-06-2021 from OpenStreetMap and setup to include primary roads only, disregarding local access roads, and including links and roads under construction.

The road graph was simplified by removing all nodes that are not intersection or dead-ends, and merging roundabouts and other junctions to a single node. After simplification 2587 nodes and 5580 edges remain. For each edge (which represents a road segment) free flow speed data was added.

For a full description of the road graph compilation including motivation, metrics and exact steps, see Appendix C: Road graph.

Digital elevation model

The digital elevation model (DEM) is the most important input variable for the stormwater model. The DEM is compiled using data from the Helsinki Spatial Data Directory (City of Helsinki, 2017). The source data contained points with XYZ coordinates in a triangular irregular network (TIN) with on average one point per 5.17 m^2.

This TIN was converted to a raster with square cells of 2 by 2 meters. Buildings, sources from the same data portal, were burned into the DEM as impenetrable structures and rivers and larger streams were carved out as sinks where water could flow away without limitation.

For a full description of the DEM processing is included in Appendix A: Digital elevation model.

Rainfall values

The second input value the stormwater model needed was a rainfall value. 6 different values were chosen. The first four are based on the 6-hour precipitation levels for 10-, 20-, 50- and 100-year return periods, the rainfall that would fall in 6 hours in a once in N-year event.

For these return periods, these are estimated at 44 mm (31-56 mm 95% CI), 54 mm (38-69), 71 mm (50-92) and 86 mm (60-112) precipitation respectively (Venäläinen et al., 2007). To construct the scenario's the best estimates were used.

Return period	**Precipitation** *Best estimate (95% CI)*
10-year	44 mm (31-56)
20-year	54 mm (38-69)
50-year	71 mm (50-92)
100-year	86 mm (60-112)

Table 3-1: Precipitation values for different return levels

The last two rainfall values were added to increase the range of the rainfall values to stretch at least one order of magnitude. The daily rainfall record in Finland is 198.4 mm, so 198 mm and 20 mm were chosen. 20 mm is less than half of the lowest 44 mm scenario, and 198 mm over double, together spanning an 10x range. These values will be used to test model validity and results for more extreme values.

Traffic data

To be able to weight each route, traffic data is needed containing the number of vehicles traveling from each point to each other point. This origin-destination data was found available from TomTom. Are few limitations did exist however.

There was a maximum number of 200 regions. This was far fewer than the 2587 intersections in the network, so not each intersection could get its own OD data.

The approach chosen was to group nearby nodes together to form a smaller number of regions. Helsinki was split in 199 regions with each region containing between the 12 and 16 intersections. A map with the defined OD regions is presented in Figure 6-20 in Appendix C.

Two datasets were generated, one included all trips that crossed through one or more of the regions, representing all traffic, and one that only included traffic that both started and ended in one of the defined regions. This way both the full impact on travel time can be measured, as well as the impact on trips within Helsinki. The full traffic dataset includes 1,911,281 trips after processing and the dataset with internal traffic only 1,398,892 trips. The usage of datasets of this size also allows for reporting very detailed statistics.

A detailed description of the traffic data used can be found in the Origin-destination data section of Appendix C.

Models & analysis

In this research two major models were used and developed: One of tracing rainfall based on elevation maps, and one for analysing the (partially flooded) road network in a computationally feasible way.

Stormwater model

The stormwater model is used to identify in which scenarios a road floods. It's based on the bluespot model developed by Balstrøm & Crawford (2018) and modified to fit this application. It used the digital elevation model to identify bluespots, which are depressions in the landscape. Each bluespot has an watershed area, which his the total area for which rainfall ends up in the bluespot. Using water stream tracing based on the elevation model, this watershed area of each bluespot was calculated. The total rainfall volume for each bluespot can be calculated using this watershed area and the amount of rainfall inputted.

From this point the model was modified to not only calculate the potential bluespot depths when fully filled, but also to calculate the depth of each point exactly when the rainfall volume is not enough to fill the bluespot fully. The toolset developed split bluespots in fully and partially filled sets and used an iterative process to estimate the depth of the water in a bluespot. This was achieved by comparing the volume of a virtual plane to the digital elevation model, calculating the volume between those, and comparing that to the rainfall volume for that bluespot.

The result was a flood map with depths for each coordinate with a 2-meter resolution, covering the full extent of the DEM. Based on the shared GK25 coordinate system between the flood map and the road network, the average and maximum water depths were added to the road segments attributes, as wel as the length of the flood, for each rainfall scenario.

This process and the exact models and parameters used are described in detail in the Determining the water level in bluespots section of Appendix B.

Road graph

With the addition of flood data from the stormwater model to each road segment, analysis on it can be performed. Two main algorithms were used, one to calculate the shortest paths between combinations of roads, and one to measure the usage of each road segment.

Dijkstra's algorithm was used to calculate the shortest paths between each combination of the 2587 nodes in the road network. From this a database could be build with the travel times and the road segments used in between for each node combination. This dataset was extended with the origin-destination traffic data, from which now weighted averages and distributions of travel times could be calculated.

The second algorithm used was betweenness centrality, which determines how many times each edge in a graph (in our case road segments in the road network) will be used by all the shortest paths in the graph.

Betweenness centrality counts each shortest path that uses an edge with the same weight: one. In our road network that isn't true, since between some points in the network there is far more traffic than between others. Therefor the implementation was modified to create a new metric: OD-weighted betweenness centrality. This metric is equivalent with the theoretical number of cars passing over each road segment if they all take the shortest route.

With these two algorithms, the two main metrics to answer our research question can be determined: How much will the average travel time increase, and which road segments were used most.

For each rainfall scenario this process had been run, with the removal of all road segments that were fully flooded in that scenario and reduced speed on partially flooded roads (discussed further in the Driving speed section of Appendix C).

The average travel time was also calculated when single road segments were removed, and the increase from the base scenario was calculated. This way the individual contribution of a road segment was calculated.

All code used for the road graph analysis can be found in the *OSMnx_road_graph_analysis_BEP.ipynb* Jupyter notebook, available [on GitHub](#).

Reductions

Modelling rainfall and waterflow over a 250 km^2 area and simulating traffic through a complete city is quite complex, hence some reductions and assumptions are made in the models to fit it in the scope of this research. With each reduction the goal was constantly to keep the models as valid as possible while keeping it computationally feasible.

Stormwater

The stormwater model included 4 major assumptions and reductions:

- Buildings have a height of 120 meters and are impenetrable. Rain falling on the roof will be added to a directly adjacent stream, adding the building area to the bluespot watershed that the stream leads to.
- The DEM only has one elevation value for each location; each XY-point only has one Z-value. This results in multi-levels structures being reduced to their highest level. This results in highway-overpasses being impenetrable structures where water can't flow under.
- For partially filled bluespots, water doesn't flow from a full upstream bluespot to a next downstream bluespot. Only its own watershed area is used to calculate it's depth.

- No water is absorbed by the sewer system or ground surfaces. Due to the short duration of the rainfall event and the high. Cities also have a high ratio of impenetrable surfaces such as concreate and asphalt, limiting water absorbed by surfaces.

Road graph

The road graph contains 3 major limitations, aside from the selection of roads and scope described in the Data assembly section earlier and in Appendix C.

- Helsinki was split into 199 regions, with each region containing between 12 and 16 nodes (intersections). Between each combination of two regions, origin-destination traffic volumes were generated. Traffic demand of each region was uniformly distributed over all the nodes in that region. Figure 6-20 in Appendix C shows the regions and corresponding nodes on a map.
- Intersections don't add travel time to pass, only the road segments determine the travel time based on their maximum speeds.
- Traffic flows freely, without congestion taking place. The travel time per road segment is only determined by the maximum speed and the level of flooding, not the amount of traffic.

The last two reductions would be representative for an emergency situation, where everyone is advised to stay indoors at much as possible. Only emergency services would use the road network, which don't have to obey traffic laws if necessary. This makes this research not applicable to situations in which heavy economic activity and its corresponding traffic take place.

Bias

Some reductions will have a distinct effect making this research more pessimistic (heavier floods or larger travel time increases), while some will make the research more optimistic (lesser floods or smaller travel time increases). This sub-section indexed those, to better understand the limits of this research. The reductions from the section above are listed in the two categories.

Aspects that make the results pessimistic:

- Multi-level structures being impenetrable.
- No water is absorbed by the sewer system or surfaces.

Aspects that make the model optimistic:

- Partially filled bluespots only use their own watershed area.
- Traffic doesn't get congested

Because the quantitative contribution of these factors on the results are unknown, it's unclear if the overall bias will be pessimistic or optimistic. But when discussing these results or doing follow up research, these are the factors to keep in mind.

Experimental setup

As described above, six different rainfall scenarios were defined and two traffic datasets were compiled. Of the former scenario's, four were based on precipitation level return periods and two on extreme events. Furthermore, two traffic scenarios were defined, one using all traffic passing through the road network, and one with only traffic with it's origin and destination in Helsinki.

For all twelve combinations of rainfall and traffic data results have been gathered. When discussing then in the Results section, the focus will be on the regular once in N-year rainfall scenario and the full traffic dataset, that includes external traffic.

For each scenario, the 12 most used road segments were also removed individually. Travel time statistics have been gathered for each of those road segments for all rainfall scenarios and both traffic datasets.

This research is fully deterministic, with no stochastic elements, nullifying the need for multiple runs of the same experiment.

4. Results

For the five sub-questions defined in the research questions, results have been gathered and will be listed and discussed in this section. First the most used roads will be discussed and then roads prone to flooding. Combining these the highest risk roads will be presented, followed by the total travel time increases for the different scenarios. Finally, the travel time increase caused by the highest risk road will be discussed.

Most used roads

To answer sub-question a), the roads most used were indexed. This provides an overview which roads are heavily used, and which are not. To find the usage of each road, origin-destination traffic data was used as weights for a betweenness centrality calculation, as described in chapter 3. The higher this OD-weighted centrality value, the more a certain road segment should be used.

Road segment	**Name**	**Type**	**Weighted centrality (ext)**	**Weighted centrality (int)**	**Used in % of trips (ext)**	**Used in % of trips (int)**
(1459, 1464)	Kehä I	trunk	77572.3	50291.0	5.545%	3.595%
(1794, 131)	Kehä I	trunk	72298.2	49862.6	5.168%	3.564%
(589, 2251)	Kehä I	trunk	72268.3	48079.6	5.166%	3.437%
(21, 20)	Tuusulanväylä	motorway	72065.4	59889.6	5.152%	4.281%
(1537, 1438)	Tuusulanväylä	motorway	67757.2	55490.7	4.844%	3.967%
(841, 21)	Tuusulanväylä	motorway	65097.7	53655.5	4.654%	3.836%

Table 4-1: Three most used road segments for each traffic dataset

For the full traffic dataset with external traffic, which contains 1,911,281 trips, road segments of the Kehä I highway were used most, with the (1459, 1464)-segment being used by up to 5.545% of all trips. The Kehä I (Ring I) is the innermost beltway of Helsinki.

Looking at internal traffic only (1,398,892 trips), the Tuusulanväylä motorway is used the most, with some segments being used around 4% of all trips. A major driver for this heavy use is the traffic between the city centre and the airport (see also the Spatial Sankey diagram in Figure 6-22). The (21, 20)-segment is used most, supporting 4.281% of all traffic, being located just north of the Pakila interchange in the Northbound direction of the Tuusulanväylä motorway.

Both road segments are displayed in Figure 4-1.

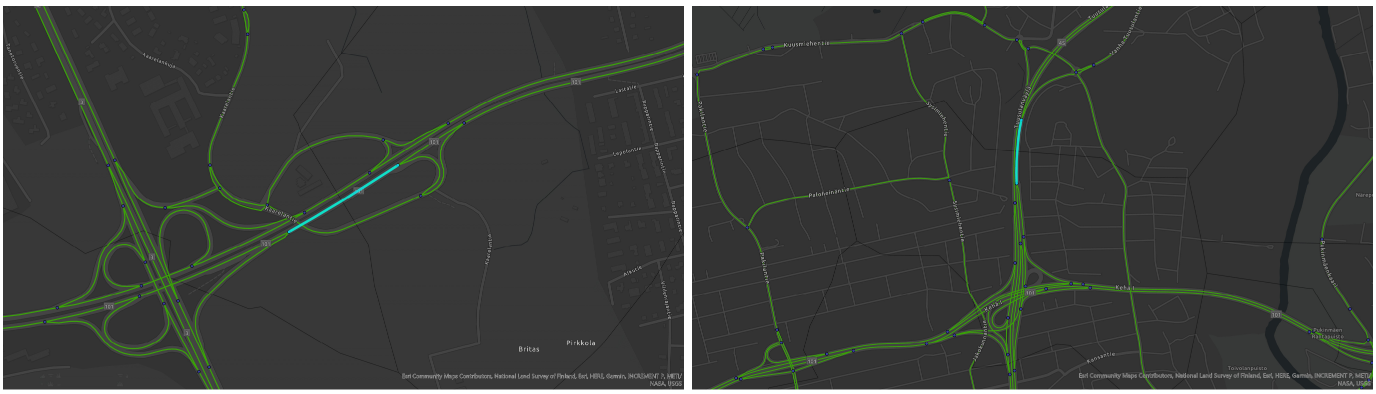


Figure 4-1: The (1459, 1464)-segment in the Kehä I (left) and the (20, 21)-segment of the Tuusulanväylä (right)

Figure 5-1 shows the OD-weighted centrality for each road segment on a map. Note the ring and spoke highways lighting up with high OD-centrality values.

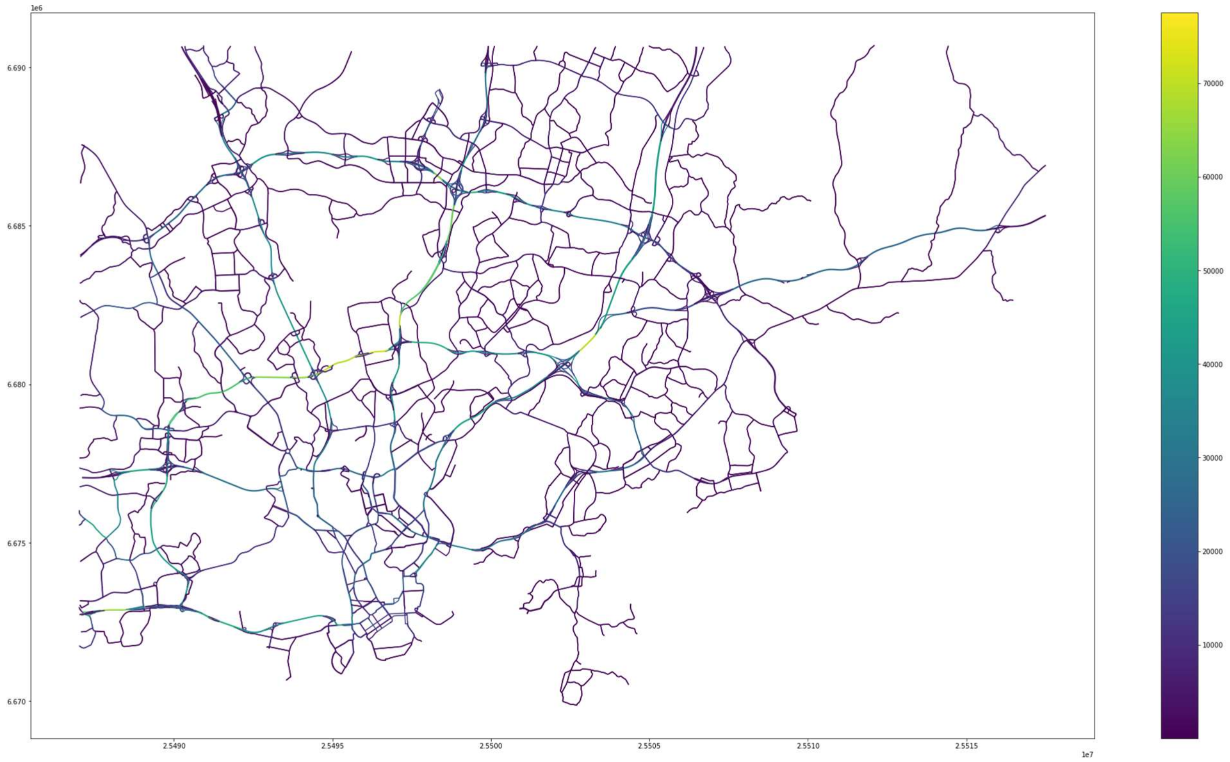


Figure 4-2: OD-weighted centrality for the OD dataset with external traffic

The roads with the highest OD-weighted centrality will be used to select roads that are removed to measure the travel time increase of individual flooded road segments for sub-question e.

Roads prone to flooding

Sub-question b, roads most prone to flooding in different rainfall scenarios were analysed. Using the digital elevation model discussed in Appendix A, and the stormwater model discussed in Appendix B, for each road the flooded length, average flood depth and maximum flood depth were determined for each of the 4 regular and 2 extreme scenarios.

Each road with a maximum flood depth exceeding 300 millimetres of water was considered flooded and inaccessible for traffic.

Rain scenario	**Additional roads flooded**	**Total roads flooded**	**Percentage roads flooded**
20	364	364	6.61%
44	122	486	8.82%
54	27	513	9.31%
71	19	532	9.66%
86	12	544	9.87%
198	22	566	10.27%

Table 4-2: Number of roads flooded per scenario

In the once in 20-year (54 mm) scenario, the Santahamna and parts of the Laajaslo island become disconnected, because the main entry road is flooded. In Figure 4-3, also note the large number of road segments flooded on the inner ring highway of Helsinki, the Kehä I, and the highways connecting the city centre district to the western and eastern peninsulas.

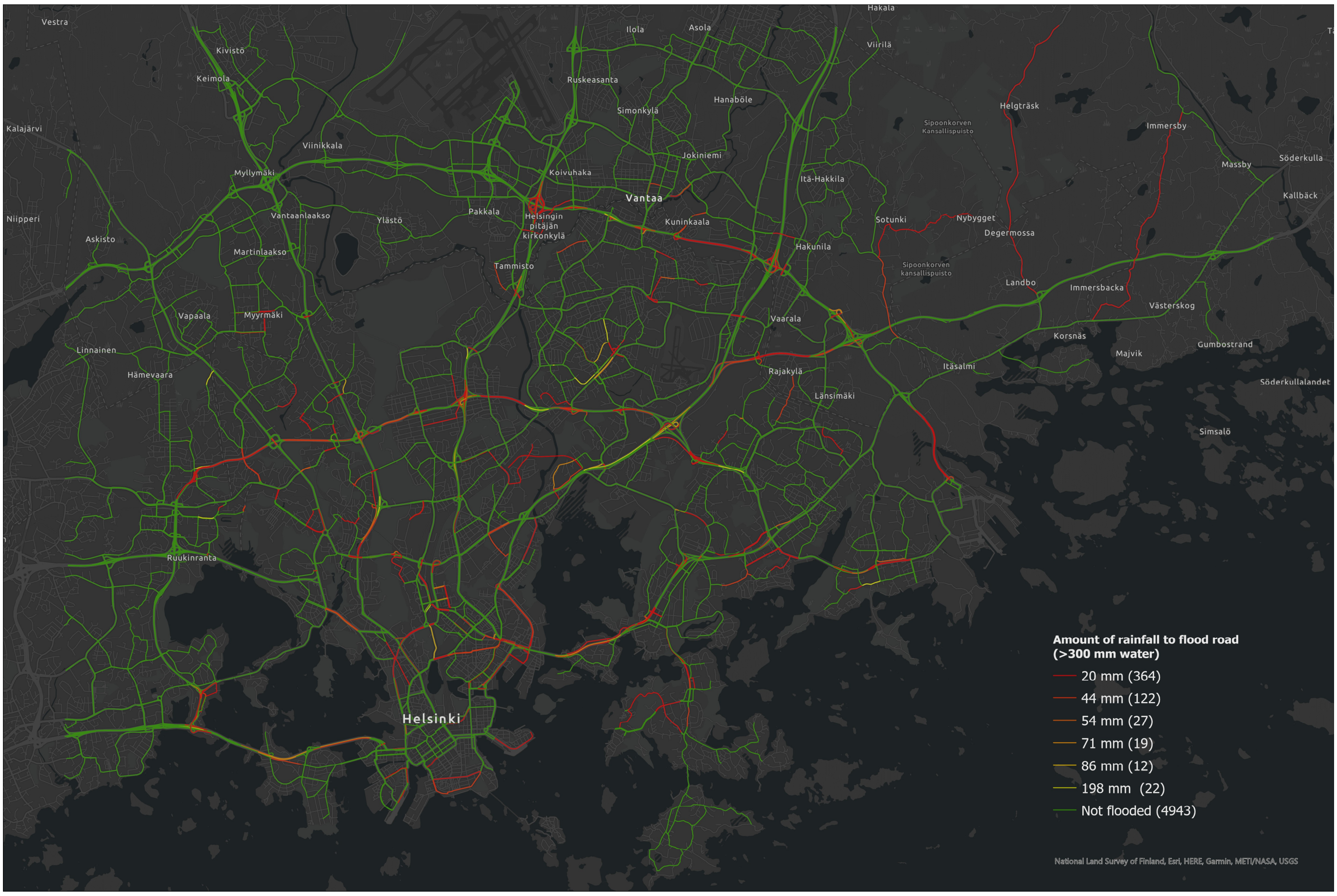


Figure 4-3: Map indicating at which rainfall scenario the road is fully flooded, with a maximum depth of more than 300 mm

The large decrease in number of available paths between nodes is most likely caused by the disconnection of the Santahamna and parts of the Laajaslo island. These decreases in paths become visible from the 44- and 54-mm rainfall scenarios. The 198 mm scenario also leads to a large decrease in available paths, mainly due to the flooding of some major highways.

Rain scenario	Available paths	Unavailable paths	% unavailable
0	6505175	187394	2.80%
44	5102003	1590566	23.77%
54	4873833	1818736	27.18%
71	4807291	1885278	28.17%
86	4747833	1944736	29.06%
20	5814948	877621	13.11%
198	4371129	2321440	34.69%

Table 4-3: Number of paths available per scenario

Selection of highest risk roads

To answer sub-question c, the results from sub-questions a and b are combined to create a selection of high-risk roads. The road segments with the highest OD weighted betweenness centrality that will be flooded in at least one of the scenarios are listed for each rainfall scenario in Table 4-4.

Road segment	Name	Type	0	44	54	71	86	20
(589, 2251)	Kehä I	trunk	72268	8	8	8	8	21733
(218, 206)	Kehä I	trunk	69787	42691	0	0	0	33710

(1463, 130)	Kehä I	trunk	68379	0	0	0	0	50185
(1777, 314)		primary_link	1963	88958	0	0	0	42364
(1543, 1777)		primary_link	1157	88889	490	485	482	42282
(1039, 1038)		trunk_link	13141	67447	0	0	0	23057
(956, 957)	Tapanilankaari	secondary	4164	38374	161556	158025	0	5484
(957, 1256)	Tapulikaupun gintie	secondary	3380	38349	161537	158005	0	5434
(1687, 35)	Lahdenväylä	motorway	34802	51987	149120	142662	2582	72419
(165, 160)	Länsiväylä	motorway	27009	46850	50054	43242	55667	47603
(49, 1040)		trunk_link	13178	66095	37704	37963	41478	39846
(1827, 1794)		trunk_link	8269	17394	27406	27189	27827	45574
(150, 153)	Länsiväylä	motorway	47983	54075	0	0	0	70194
(2194, 2514)	Länsiväylä	motorway	43677	45730	7727	7379	0	63093

Table 4-4: Most used roads that flood in at least one rainfall scenario (with the full traffic dataset)

Notable is the lack of large changes between the 54 mm and 71 mm scenarios. This mean no additional high-centrality road segments are flooded in the 71 mm scenario compared to the roads already flooded with 54 mm rainfall.

Two notable combinations of two high-centrality edges that eventually will be flooded are the (1543, 1777)-(1777, 314)-combination and the (956, 957)-(957, 1256)-combination. The former are two segments of the type of primary_link, meaning highway entrances, exits or transfers. In this case, it's a highway entrance and exit that are used to avoid a flood on the (1543, 314)-segment of the Hakamäentie, the Finnish regional road 100. Such a turn wouldn't be possible, so this is a model limitation.

The second combination is a secondary road that is used significantly starting from the 44 mm scenario, and it the most used road in the 54- and 71-mm scenarios. The secondary road becomes the fastest route after segments on the Tuusulanväylä and Kehä III highways flood. It stops being used in the 86 mm scenario, when both segments are flooded.

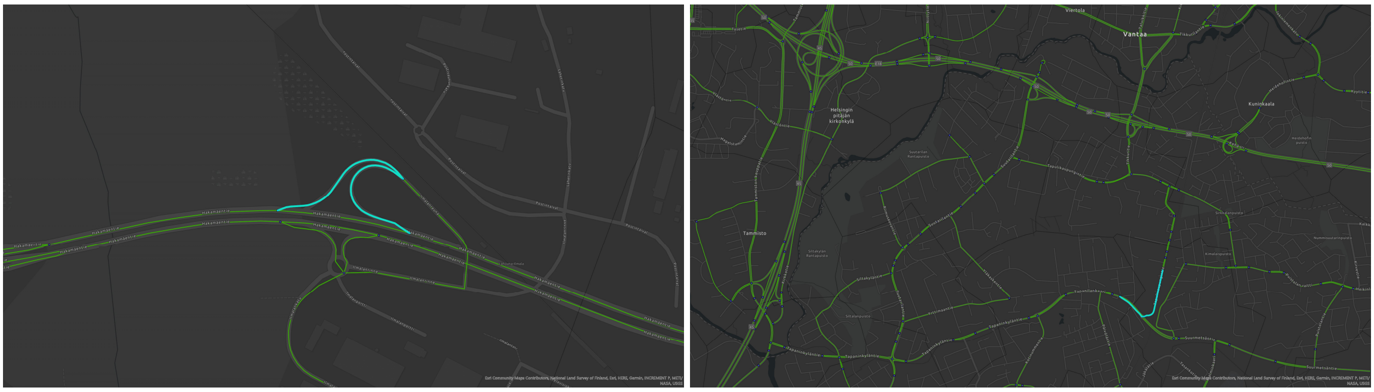


Figure 4-4: The (1543, 1777)-(1777, 314)-combination (left) and the (956, 957)-(957, 1256)-combination (right)

The full data can be found on GitHub in the *proc_central_flooded_edges-12.xlsx* spreadsheet, which lists the 12 most used roads in each flood scenario for both the external traffic and internal traffic only datasets.

Total travel time increase in different scenarios

To answer sub-question d, the total travel time in the different scenarios is compared. In the base scenario without rain, a trip in external traffic dataset takes 491.5 seconds on average, and with internal traffic only 435.7 seconds. In each rainfall scenario this increases significantly, but can't be directly compared with that base scenario, because the set of trips that is still possible changes. Longer trips have a higher chance to be impossible due the higher amount of road segments it uses, and thus a higher change that a least one of them is flooded.

To compensate for this, the mean travel time (MTT) for the base scenario without floods is also calculated using only the trips that are possible in the flooded scenarios it's compared against. This value is displayed in the *MTT on id. set without flood* column in Table 5-1 and 5-2 and used for the MTT difference and increase calculations.

Rain (mm)	**Trips possible**	**Total travel time (s)**	**Mean travel time (s)**	**MTT on id. set without flood (s)**	**MTT difference (s)**	**MTT increase**
0	1.911E+06	9.394E+08	491.511	491.5108	0	0.0%
44	1.631E+06	1.156E+09	708.865	486.3331	222.531	45.8%
54	1.596E+06	1.392E+09	872.525	486.1419	386.3828	79.5%
71	1.583E+06	1.394E+09	880.499	486.1625	394.3363	81.1%
86	1.569E+06	1.672E+09	1065.576	486.2918	579.2842	119.1%
20	1.768E+06	1.085E+09	613.894	488.8385	125.0551	25.6%
198	1.487E+06	1.613E+09	1084.374	486.4185	597.9557	122.9%

Table 4-5: Mean travel times in different scenarios with external traffic

For the traffic dataset with external traffic, the mean travel time on the identical set increases with 45.8%, 79.5%, 81.1% and 119.1% for the once in 10-, 20-, 50-, and 100-year scenarios respectively. In the extreme scenarios, 20 mm of rain increases the travel time by 25.6% and 122.9%.

Notable is the only very small increase between the once in 20- and 50-year scenarios. A possible explanation is that no major highways are additionally flooded with 71 mm rain compared to 54 mm, leading to only a very minor 1.6 percent-point increase in mean travel time. For the 86 mm to the extreme 198 mm rainfall scenarios the same possible explanation can be used. Note however, that in the 198 mm scenario a lot of additional trips become impossible.

Rain (mm)	**Trips possible**	**Total travel time (s)**	**Mean travel time (s)**	**MTT on id. set without flood (s)**	**MTT difference (s)**	**MTT increase**
0	1.399E+06	6.095E+08	435.694	435.694	0	0.0%
44	1.153E+06	7.548E+08	654.342	426.685	227.657	53.4%
54	1.124E+06	9.095E+08	809.091	425.944	383.147	90.0%
71	1.113E+06	9.093E+08	817.070	425.847	391.223	91.9%
86	1.102E+06	1.147E+09	1040.844	425.774	615.070	144.5%
20	1.273E+06	7.201E+08	565.815	431.203	134.612	31.2%
198	1.032E+06	1.093E+09	1059.496	424.179	635.316	149.8%

Table 4-6: Mean travel times in different scenarios with internal traffic only

Looking at internal traffic only, the same pattern is visible: Major increases in travel time up to 54 mm of rain, a small increase with 71 mm, and then a large increase to 86 mm, with the extreme 198 mm of rain not adding much additional travel time. The increase in travel time to a limited extent larger than when the external travel is included, but generally comparable.

Travel time impact of high-risk road segments

To answer sub-question e, the highest risk roads are removed from the graph one by one and for each new network the travel times were calculated for both the external traffic dataset, the last two for the internal traffic only.

In Table 4-7 the 3 road segments that lead to the highest increase in travel times are reported. The table contains a few missing values due to the roads falling within the 12 highest risk for one traffic dataset, but not for the other.

The *Flooded at* column contains the rainfall scenario at which a road starts flooding to an amount that becomes undrivable (>0.30 m). For example, a road segment that floods at 44 mm, also floods in the higher rainfall scenarios.

Road segment	**Name**	**Type**	**Flooded at (mm)**	**Ext. delay (s)**	**Ext. increase**	**Int. delay (s)**	**Int. increase**
(1762, 1851)	['Kehä I', 'Mestarintunneli']	trunk	20	2.846	0.579%		
(2030, 220)	Kehä I	trunk	20			2.373	0.545%
(324, 1686)	Sörnäisten rantatie	primary	20			1.842	0.423%
(1463, 130)	Kehä I	trunk	44	5.011	1.019%	4.819	1.106%
(214, 1409)	Kehä I	trunk	44	4.101	0.834%	3.548	0.814%
(18, 1977)	Kehä I	trunk	44	3.526	0.718%		
(150, 153)	Länsiväylä	motorway	54	8.215	1.672%	5.821	1.337%
(218, 206)	Kehä I	trunk	54	4.844	0.985%	4.459	1.023%
(127, 128)	Kehä I	trunk	54	3.634	0.739%	3.347	0.768%
(149, 150)	Länsiväylä	motorway	71	1.416	0.288%		
(146, 2515)	Länsiväylä	motorway	71	0.996	0.203%	0.735	0.169%
(918, 919)	Kotinummentie	secondary	71	0.026	0.005%	0.033	0.008%
(2194, 2514)	Länsiväylä	motorway	86	1.994	0.406%	1.469	0.337%
(1409, 540)	Kehä I	trunk	86	0.873	0.178%	0.665	0.153%
(381, 331)	Pihlajamäentie	secondary	86	0.365	0.074%	0.466	0.107%
(589, 2251)	Kehä I	trunk	198	4.429	0.901%	4.029	0.925%
(1687, 35)	Lahdenväylä	motorway	198	2.056	0.418%	1.445	0.332%
(1742, 841)		motorway_link	198	1.904	0.387%	1.904	0.437%

Table 4-7: Travel time increases caused by individual road segments

At the 44 mm scenario, road segments of the Kehä I (Ring 1), the inner-most highway ring of the Greater Helsinki region, are flooded. They lead to large increases in travel time, with the (1463, 130) segment, just West of Länsi-Pakila, increasing average travel times with 1.019% and 1.106% for the full and internal only traffic dataset respectively.

With 54 mm of rain the (150, 153)-segment on the Länsiväylä (Western Highway) starts to flood. This motorway connects the centre district with the western parts of Helsinki, including the Lauttasaari island. Notable is that traffic starting or ending outside Helsinki is affected most, with a travel time increase of 1.672% for the full dataset. When looking at internal traffic only, the delay is a smaller but still very significant 1.337% increase in average travel time.

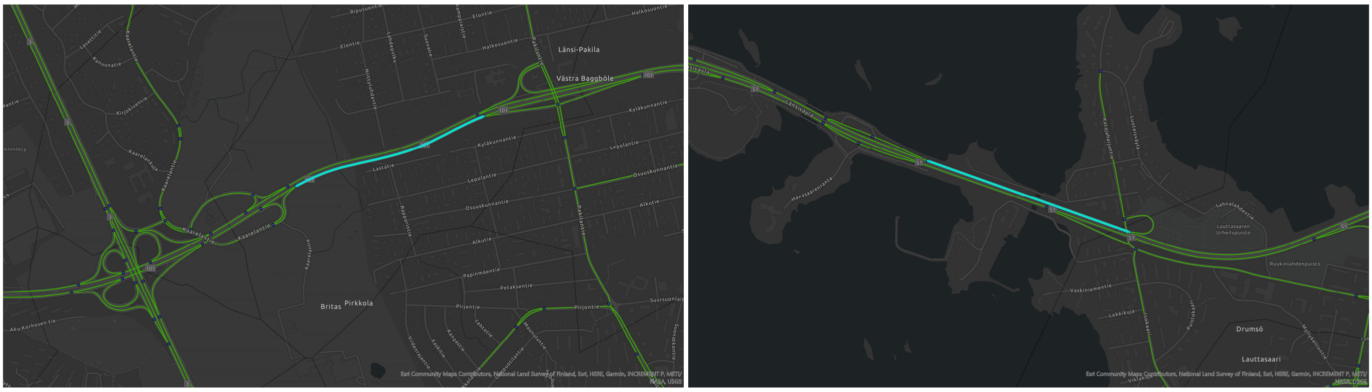


Figure 4-5: The (1463, 130)-segment in the Kehä I (left) and the (150, 153)-segment of the Länsiväylä (right)

In the 71 mm scenario additional segments of the Länsiväylä motorway flood, but those segments have a way smaller impact since there are alternative interior routes possible over the Lauttasaari island. Most likely those interior, secondary roads can't handle the full traffic of a motorway, so the slowdown would be larger in heavy traffic.

The 86 mm scenario continues to flood segments of the Länsiväylä and Kehä I highways. These segments on their own also introduce significant slowdowns, but not as large as some of the segments that started flooding at 44 and 54 mm.

Looking at the extreme scenario's, at 20 mm the (1762, 1851)-segment Mestarintunneli, a tunnel part of the western Kehä I, already starts flooding, leading to 0.579% increase in average travel time for the traffic dataset with external trips, but not making the top 12 travel time increases for internal traffic.

At 198 mm, the (589, 2251)-segment floods, another segment on the northern part of the Kehä I, leading to a 0.901% and 0.925% increase in average travel time for the full and internal only traffic datasets respectively.

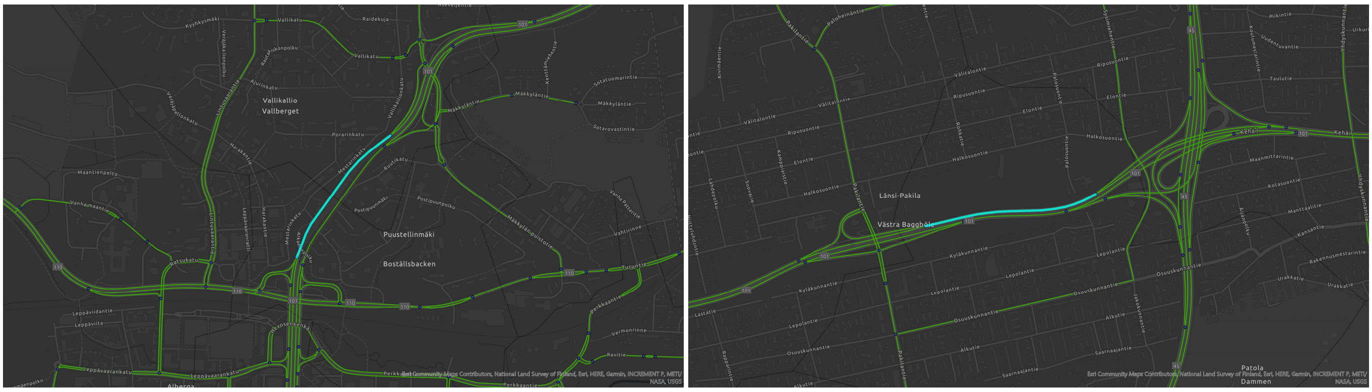


Figure 4-6: The (1762, 1851)-segment in the Mestarintunneli (left) and the (589, 2251)-segment of the Kehä I (right)

An extended version of Table 4-7 can be found in Appendix E, Table 6-5, containing more road segments per scenario that cause a large increase in average travel times. The full data is available on GitHub in the *proc-hr-edges-stats.xlsx* spreadsheet.

5. Conclusion and discussion

This section will discuss the results and formulate conclusions, relate them to approach and the limitations of this study, and discuss future research.

Conclusion

A few major highways contribute most to the travel time increases. The Kehä I highway and Tuusulanväylä motorway contribute very significantly to these travel time increases. Multiple individual segments on these roads, as well as on the Länsiväylä motorway, contributed to average travel time increases larger than one percent for the whole Helsinki road network.

The mean travel time increases of 45.8%, 79.5%, 81.1% and 119.1% using the full traffic dataset for the once in 10-, 20-, 50-, and 100-year rainfall scenarios are quite significant. It's notable that the increase of rainfall from 54 mm to 71 mm for the once in 20- and 50-year scenarios doesn't lead to any additional major highways flooding, and thus also a very low increase in average travel time.

Looking at internal traffic only, which could be a better estimation of actual traffic demand in an extreme weather event, mean travel time increases of 53.4%, 90.0%, 91.9% and 144.5% were measured for the regular rainfall scenario's. This is a small but significant increase compared to the full traffic dataset, especially in the once in 100-year scenario with 86 mm rainfall, in which more parts of the Länsiväylä and Kehä I flood.

The extreme scenario's show that 20 mm rain already offers a within this model 25.6% and 31.2% increase in average travel time, which is most likely a overestimation, since the drainage system will absorb this relatively low amount of water well. The 198 mm rainfall scenario doesn't see large increases in average travel time, 122.9% and 149.8% for both traffic datasets respectively, most likely due to no major additional road segments flooding.

Discussion

Two major reductions limit the validity of this research. The simplification of multi-level structures being represented as solid features created bluespots near highway over- and underpasses that shouldn't be there, a pessimistic effect. The downstream bluespots not being filled by overflowing upstream bluespots meanwhile makes some bluespots less deep than they should be, an optimistic effect.

Furthermore, the scope of this research is somewhat limited by the assumptions in the road graph, mainly to the disregard of traffic congestion that will most likely take place when major highways are offloaded to secondary roads. While this can be representative for emergency vehicles, it won't be for regular traffic.

Less important but still notable, while the traffic datasets are quite large with over 1 million trips each, which allows for reporting detailed statistics with many significant figures, they only reflect 4 weeks of traffic in January 2020 (before the COVID-19-pandamic), which will vary seasonally. Traffic demand will also influenced by extreme weather events.

Future research

In future research, significant improvements to the stormwater model could be made by implementing multi-level structures such as highway overpasses and including the remaining watershed volume of fully filled upstream bluespots when filling downstream bluespots. This will improve the accuracy of the flood maps, on which roads will flood for which rainfall values

Traffic data could be improved by increasing the resolution, by using more and smaller regions or even provide data on intersection level, and adding more different scenarios, for example rush hour or holiday traffic.

The road graph could be improved by implementing reduced speed based on traffic volumes and representing intersections in more detail. To do this, an increase in computational power and/or efficiency is also needed, which was one of the limitations in this research.

6. References

Baky, M. A. A., Islam, M., & Paul, S. (2019). Flood Hazard, Vulnerability and Risk Assessment for Different Land Use Classes Using a Flow Model. Earth Systems and Environment, 4(1), 225–244. https://doi.org/10.1007/s41748-019-00141-w

Balstrøm, T., & Crawford, D. (2018). Arc-Malstrøm: A 1D hydrologic screening method for stormwater assessments based on geometric networks. Computers & Geosciences, 116, 64–73. https://doi.org/10.1016/j.cageo.2018.04.010

Chen, N., Yao, S., Wang, C., & Du, W. (2019). A Method for Urban Flood Risk Assessment and Zoning Considering Road Environments and Terrain. Sustainability, 11(10), 2734. https://doi.org/10.3390/su11102734

City of Helsinki Executive Office. (2021, March 25). Traffic volumes in Helsinki. Helsinki Region Infoshare. https://hri.fi/data/en_GB/dataset/liikennemaarat-helsingissa

Climateguide. (2007). Precipitation return levels in Finland. Ilmasto-Opas.Fi. https://ilmasto-opas.fi/en/datat/sateiden-toistuvuustasot

Das, D., Ojha, A. K., Kramsapi, H., Baruah, P. P., & Dutta, M. K. (2019). Road network analysis of Guwahati city using GIS. SN Applied Sciences, 1(8), 906. https://doi.org/10.1007/s42452-019-0907-4

Deloitte & Touche LLP, Curtis, P., & Carey, M. (2012, October). Risk assessment in practice. Committee of Sponsoring Organizations of the Treadway Commission. https://www2.deloitte.com/mn/en/pages/governance-risk-and-compliance/articles/risk-assessment-in-practice.html

Du, S., Shi, P., Van Rompaey, A., & Wen, J. (2015). Quantifying the impact of impervious surface location on flood peak discharge in urban areas. Natural Hazards, 76(3), 1457–1471. https://doi.org/10.1007/s11069-014-1463-2

Džebo, S., Šarić, A., Reeves, S., Ljevo, A., & Hadžić, E. (2019). Flood Impact and Risk Assessment on the Road Infrastructure in Federation of Bosnia and Herzegovina. Advanced Technologies, Systems, and Applications IV -Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT 2019), 276–289. https://doi.org/10.1007/978-3-030-24986-1_22

El Morjani, Z. E. A., Seif Ennasr, M., Elmouden, A., Idbraim, S., Bouaakaz, B., & Saad, A. (2016). Flood Hazard Mapping and Modeling Using GIS Applied to the Souss River Watershed. The Handbook of Environmental Chemistry, 57–93. https://doi.org/10.1007/698_2016_69

Environment.fi. (02–11-20). Environment > Flood mapping. https://www.environment.fi/floodmaps

ESRI. (2021). 2D, 3D & 4D GIS Mapping Software | ArcGIS Pro. Retrieved March 29, 2021, from https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Finnish Meteorological Institute. (2021). Open data sets. https://en.ilmatieteenlaitos.fi/open-data-sets-available

HRI. (2019, December 2). Helsinki Region Infoshare. Open Data Service. https://hri.fi/en_gb/

Juga, I., Hippi, M., Nurmi, P., & Karsisto, V. (2014). Weather factors triggering the massive car crashes on 3 February 2012 in the Helsinki metropolitan area. Conference: 17th International Road Weather Conference La Massana Andorra, 30th, 21, 1. https://www.researchgate.net/publication/260025554_ID_0021_Weather_factors_triggering_the_massive_car_crashes_on_3_February_2012_in_the_Helsinki_metropolitan_area

Pregnolato, M., Ford, A., Wilkinson, S. M., & Dawson, R. J. (2017). The impact of flooding on road transport: A depth-disruption function. Transportation Research Part D: Transport and Environment, 55, 67–81. https://doi.org/10.1016/j.trd.2017.06.020

Rincón, D., Khan, U., & Armenakis, C. (2018). Flood Risk Mapping Using GIS and Multi-Criteria Analysis: A Greater Toronto Area Case Study. Geosciences, 8(8), 275. https://doi.org/10.3390/geosciences8080275

Ruosteenoja, K., Jylhä, K & Kämäräinen, M. 2016. Climate projections for Finland under the RCP forcing scenarios. Geophysica, Volume 51, Issue 1: 17–50. http://www.geophysica.fi/pdf/geophysica_2016_51_1-2_017_ruosteenoja.pdf

ScienceDirect. (2021). Betweenness Centrality - an overview. ScienceDirect Topics. https://www.sciencedirect.com/topics/computer-science/betweenness-centrality

Venäläinen, A., Saku, S., Kilpeläinen, T., Jylhä, K., Tuomenvirta, H., Vajda, A., Räisänen, J. & Ruosteenoja, K. 2007. Sään ääri-ilmiöistä Suomessa. (Aspects about climate extremes in Finland. Abstract in English.) Ilmatieteen laitos, Helsinki. Ilmatieteen laitoksen raportteja 2007:4. 81 s. http://hdl.handle.net/10138/1138

Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., & Käyhkö, J. (2010). National scale assessment of climate change impacts on flooding in Finland. Journal of Hydrology, 391(3–4), 333–350. https://doi.org/10.1016/j.jhydrol.2010.07.035

Appendix A: Digital elevation model

One of the inputs needed for the stormwater model is a Digital Elevation Model (DEM). Such a model precisely describes the elevation of a corresponding terrain or area.

Selection

The highest resolution DEM found publicly available was published by the City of Helsinki (2019). Other DEMs were also available, including the EU-DEM (European Environment Agency, 2017), the ArcticDEM (Polar Geospatial Center, 2021) and NASADEM (NASA, 2020).

DEM	**EU-DEM**	**ArcticDEM**	**NASADEM**	**Helsinki DEM**
XY-resolution	25x25 meter	2x2 meter	30x30 meter	Triangular and 1x1 meter
Extend	Most EEA member countries	Arctic region, 60th North parallel and up	Global	Wider City of Helsinki

Table 6-1: Overview of publicly available DEMs

The EU-DEM and NASADEM were not suitable due to their low resolution, which wouldn't allow for accurate water flow tracing and separating roads close to each other. The most southern point included in this research had a latitude of 60.14 degrees North, which made the ArcticDEM a viable option. The Helsinki DEM was finally selected due considering it provided the highest resolution and was maintained by the local government themselves.

The The full Python script to download all the elevation tiles and write them to a single CSV file can be found on GitHub (ter Hoeven, 2021).

Data processing

The XYZ point data was subsequently processed to data format compatible with the Stormwater model. From the CSV file all the rows were loaded in as an XYZ point in ArcGIS. From this collection of points, a Triangular Irregular Network (TIN) layer was created. This TIN was delineated to contain no lines between two points longer than 20 meters.

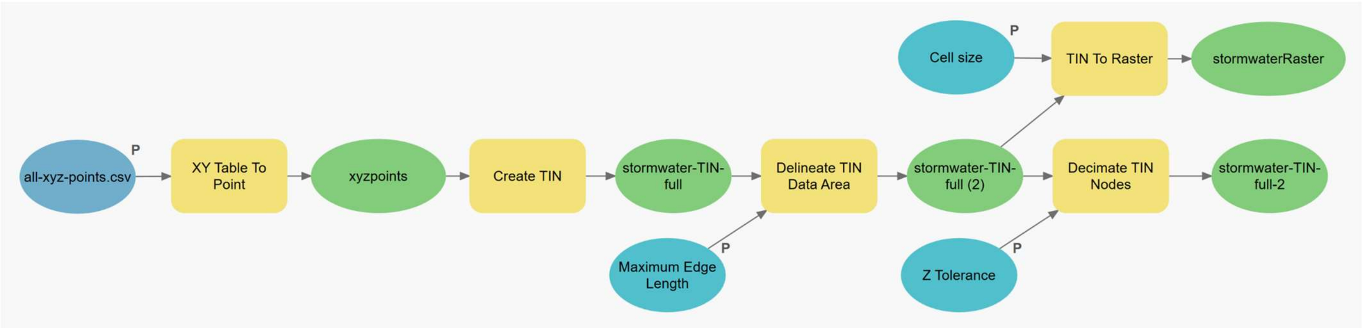


Figure 6-1: ArcGIS Pro model for processing XYZ points to TIN and Raster formats

The TIN covered an area of 271.88 km^2 and consisted of 52.515 million points, averaging a resolution of 0.193 points per m^2, or one point per 5.17 m^2. Because the irregularity of triangular data, some areas contained a higher point density then others.

Considering the requirement to be able to separate parallel roads, entrances and exits and the different driving directions, it was determined at least an resolution. Because a lot of noise can occur when sampling from a single or low number of points, including moiré and aliasing, higher resolution was preferred.

Higher resolutions would increase the workload significantly however, a doubling in resolution would result in four time the amount of data, a quadratic relation. This increased both memory and computing (time) requirements.

A resolution too high could also imply a false sense accuracy by implying data that is not actually there. A TIN is vector based, so can be exported in as high resolution as required but is still based on distinct points. All points in between are linearly interpolated.

A balance was struck on a XY-resolution of 2 meters. This resulted in a 12500 by 9500 pixel grid of 118.75 megapixels, spanning 25 by 19 kilometres. The resolution is high enough to accurately model and distinct roads, while keeping processing and memory requirements withing acceptable limits. The raster was saved in TIF format with a signed 32-bit floating point value for each pixel height. All pixels without height information from the TIN got a NoData value.

Additionally, the model generated a smaller TIN-file which described the elevation with a minimum accuracy (Z Tolerance). For a Z tolerance of 0.05 meters 34.82 million points remained, and for a Z tolerance of 0.2 meter only 12.59 million points were needed.

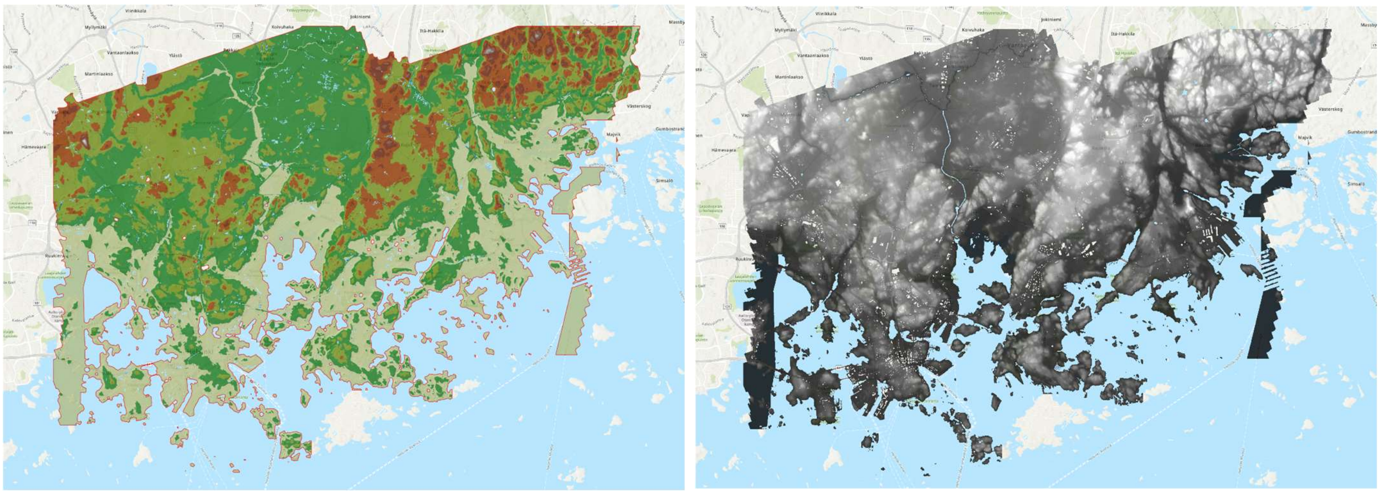


Figure 6-2: The TIN (left) and raster (right) representations of the Helsinki DEM (in ArcGIS)

Buildings and rivers

To prepare the raster DEM for processing. A few extra steps are needed. Currently, buildings are not present, and rivers and streams are filled up in some cases. The buildings need to be burned in a solid objects, and rivers need to be cut out as draining points where water can flow away. Due to the short duration of the rainfall, it's assumed rivers are not limited in their capacity to take in water (see reductions and limitations in chapter 3).

To start with the buildings, a dataset was selected and downloaded from the Helsinki Map Service (City of Helsinki, 2018). The *Buildings in Helsinki* dataset was selected due to containing all buildings also present in their basemaps and being available in polygon (vector) format.

The area within closed building complexed is filled up, to prevent huge bluespots piling up inside building complexes, which isn't relevant for this research.

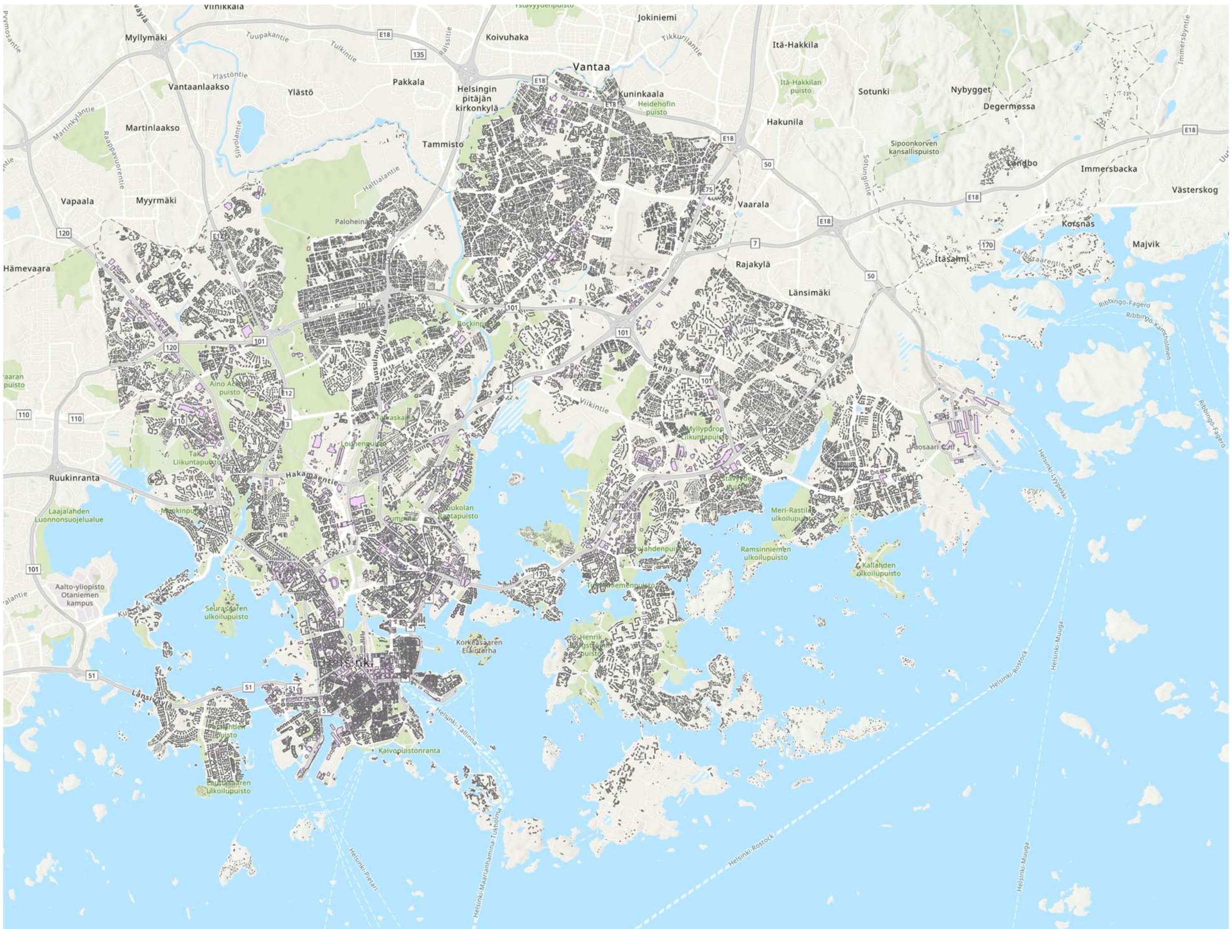


Figure 6-3: The building areas to be cut out the stormwater raster

The water areas were combined from multiple sources. The 3 following layers were used from the *Land and water areas of Helsinki* (City of Helsinki, 2020) and downloaded from the Helsinki WFS server.

- *Seutukartta_maankaytto_joet* (rivers)
- *Seutukartta_maankaytto_jarvet* (lakes)
- *Seutukartta_maankaytto_merialue* (sea area)
- *LTJ_avoin_vesi_purot* (open water streams)

The open water streams were line feature, to convert them to polygons that could be cut out of a raster a radius of 2 meters was buffered around them. All the water features were merged into a single feature layer. That feature layers was clipped to the minimum bound of the stormwater raster, to reduce the number of unneeded features. Then boundaries were dissolved to fuse connected features, again reducing the number of features.

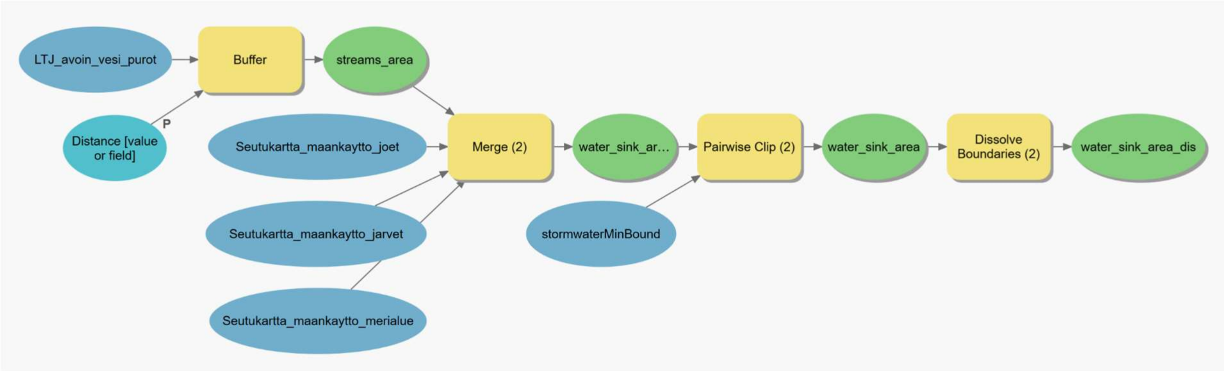


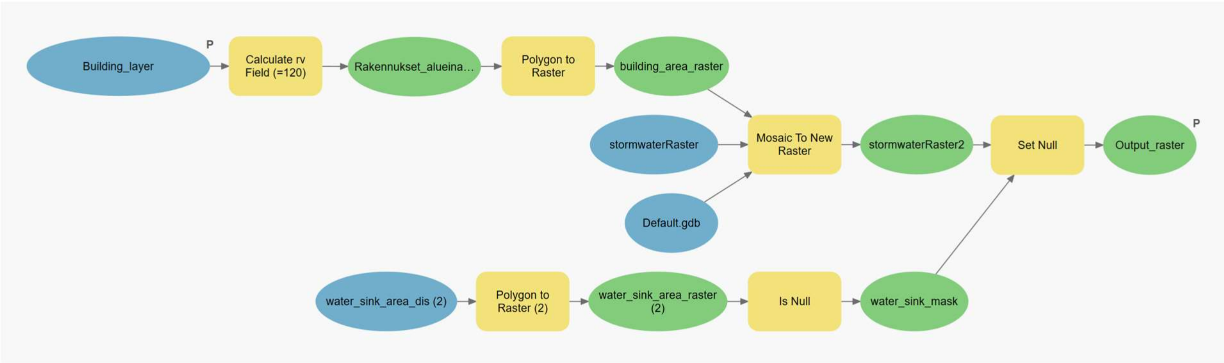
Figure 6-4: The model preparing the water polygon feature

Figure 6-5: Model to determine the minimum polygon containing the stormwater raster

Figure 6-6: The final water area to be cut out the stormwater raster

Final DEM

To complete to stormwater raster to be used as DEM, both the building and water polygon layers are converted to rasters (aligned and with the same cellsize as the stormwater raster). All the buildings are merged in with a fixed height of 120 meters (10 meters higher than the highest point in the base raster), more than enough to make them act like solid objects. Finally, the water area raster is inverted and compared with the stormwater raster, where the areas without water are kept.



The final DEM is a 12231 by 9499 pixel raster with a cell size of 2 by 2 meters. Each cell (or pixel) has a height value between -11.48 meters and 120 meters. The elevation when excluding the burned in buildings is 90.60 meters. The total area covered with elevation values is 254.59 km^2 (which excludes water areas). It covers the Helsinki municipality fully and covers small neighbouring parts of the Vantaa and Espoo municipalities.

Figure 6-7: Model burning buildings into and removing water areas from the stormwater raster

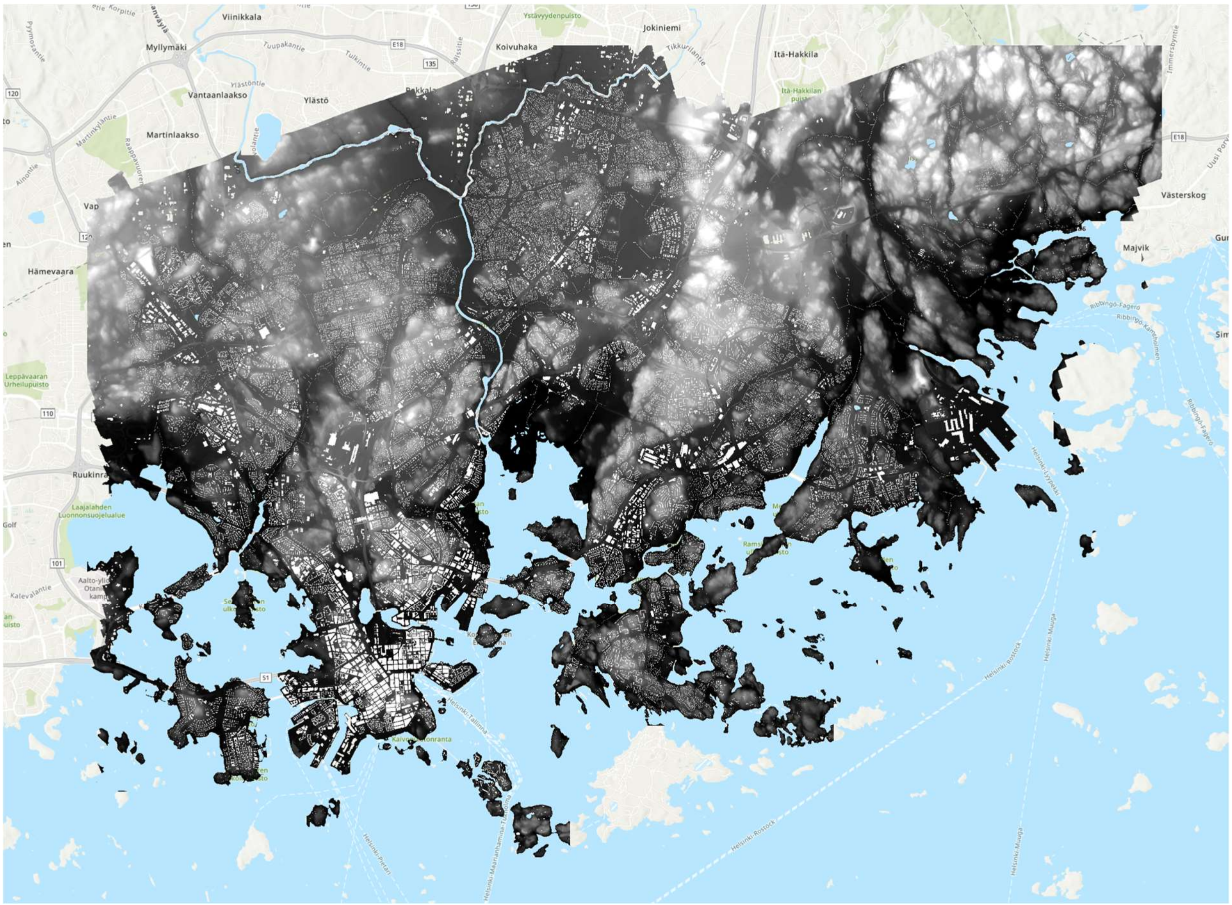


Figure 6-8: The final stormwater raster with buildings burned in (white) and rivers cut out (blue on basemap)

Appendix B: Stormwater model

The stormwater model determines where rainfall ends up. It assumes an equal amount of rainfall over the whole raster and no drainage of water. It uses the Digital Elevation Model described in Appendix B as input, together with a rainfall value (in millimetres).

Bluespot model

The first step is determining areas that could function as a bluespot, a depression in the landscape that could function as rainwater reservoir. To detect these bluespots, a currently proprietary toolbox from Thomas Balstrøm was used. The methods used are described in the research of Balstrøm & Crawford (2018). In short, local sinks (that are deeper than the tolerance level) are identified, the locations of their pour point are detected and their capacities are calculated. Then flow paths from each pourpoint to the next (local) sink are defined.

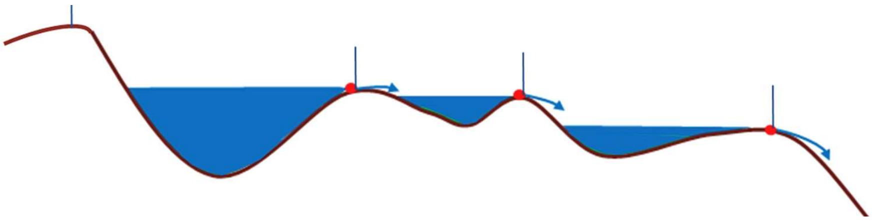


Figure 6-9: 2D cross-section of a landscape with sinks (blue) supplied with runoff from their local catchments (delineated at the vertical bars) and from upstream sinks spilling over at their pour points (red). Illustration from Balstrøm & Crawford (2018)

Running this model on the stormwater raster described in Appendix B, resulted in 174587 detected bluespots. Each bluespot has a polygon shape, and attributes for the capacity, maximum depth and watershed area. A map with all bluespots and their depth is included in Figure 8-10.

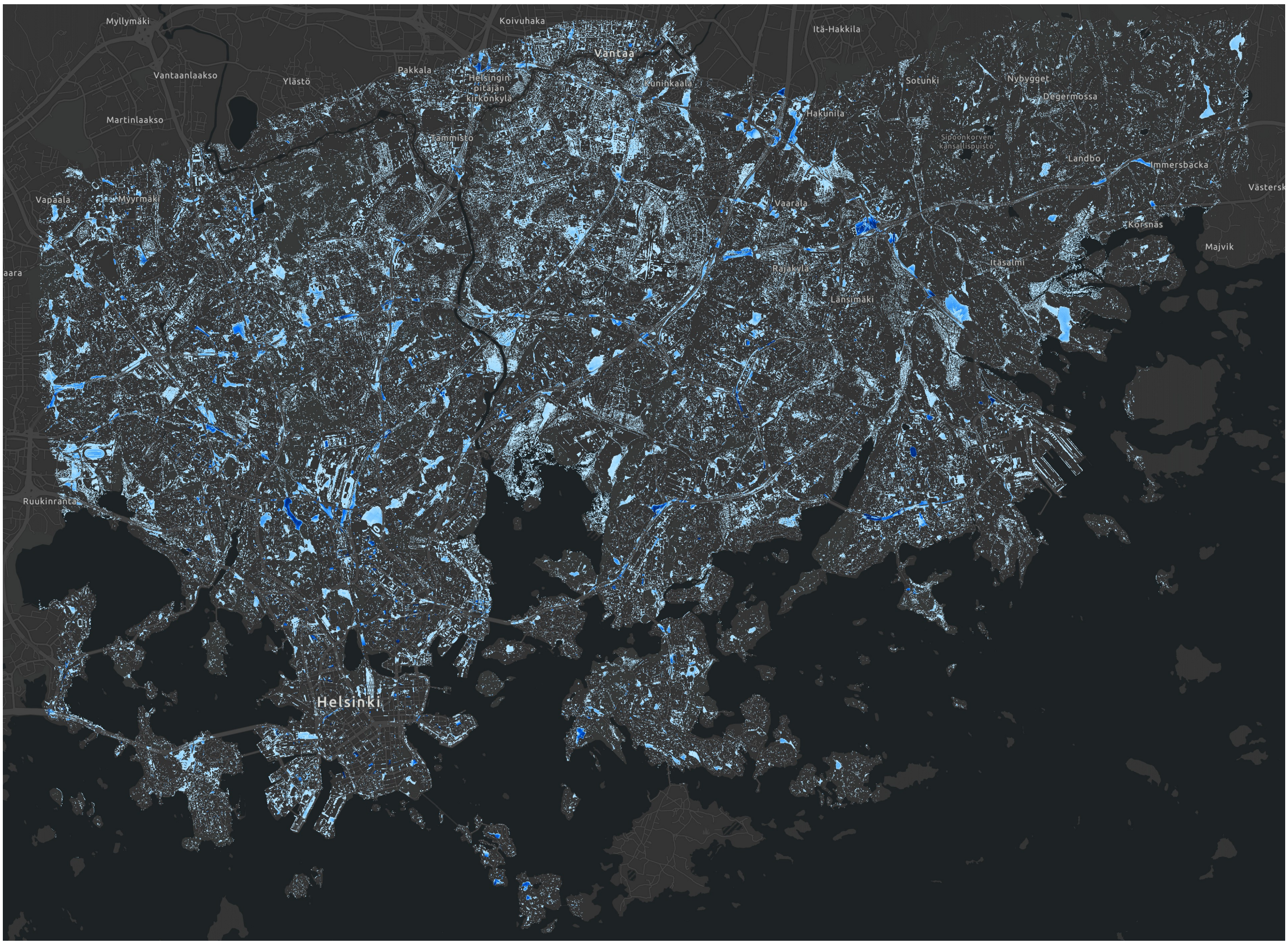


Figure 6-10: Map of all bluespots. Darker blue indicates a higher bluespot depth.

Each bluespot will collect a different amount of water from its watershed area (the area from which water flows into that bluespot) in each rainfall scenario, and thus be filled to a different level. These water levels for each scenario need to be calculated. The method chosen and implemented approaches the water depth based on the rainfall volume (the water collected from the watershed area), the capacity and the area. As the bluespot fills up, the area increases, however.

Due to this extensive amount of processing, all bluespots that even when fully filled do not intersect a road, are not relevant for this research. Therefore, these bluespots are removed from the selection, keeping only 4204 of the 174587 bluespots, 2,41%. This reduces the amount of processing that is needed to be done by over a factor 40.

Since all bluespots have a certain capacity and will be filled to a different level with different amount of rain, removing these bluespots will have impact on the level the remaining bluespots is filled. However, since all bluespots have already their watershed area calculated before removing the unneeded bluespots, this is an acceptable reduction considering the reduction in processing time.

Determining the water level in bluespots

With the identified bluespots and their properties a model was built to determine the exact depth of each bluespot. The model consisted of 3 main tools, with one using 3 other tools. All tools were developed in and tested with ArcGIS Pro 2.8.

1. First, a subset of the bluespots needs to be selected. The *Select Bluespots Create Raster and TIN* tool selects a portion of the *Bluespots* that intersect with the *Edges* (roads in this case) and

saves it as a new feature *BluespotsS*. Then it filters the *BSDepths* raster to only have values if it intersects with *BluespotsS*, save it as *BluespotsRaster2.* Finally, it converts this raster and to a triangular irregular network (TIN) layer called *BluespotsTIN*, so that depth measurement can be done from it. converts it to a TIN called *BluespotsTIN.* This TIN layer represents a depth map of each bluespot, so that the area and volume can be calculated from it.

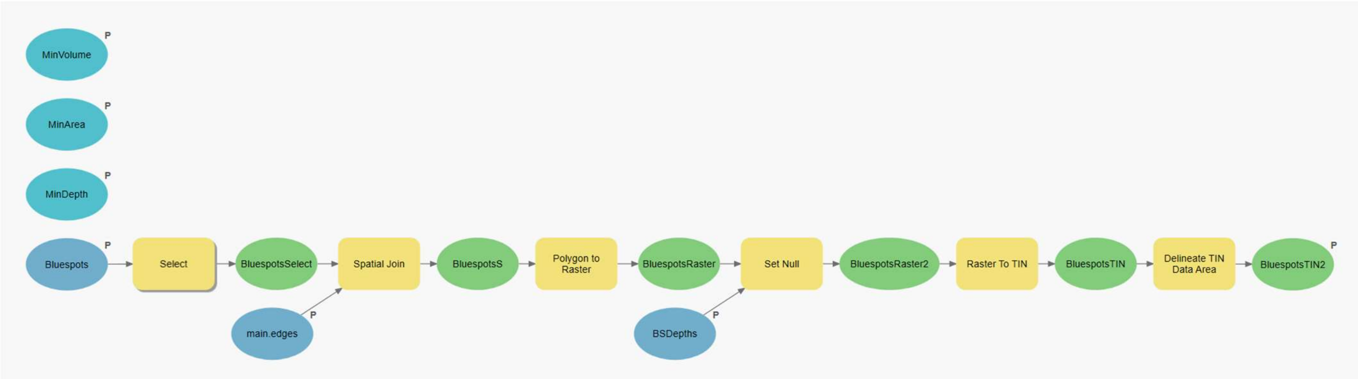


Figure 6-11: Select Bluespots Create Raster and TIN model

2. The *Bluespots Depths Assembled 2* tool will take the selected bluespots, a level of rain (in millimetres) and the reference TIN as inputs and calculate the actual depth at point of each bluespot. It consists of 3 separate tools:

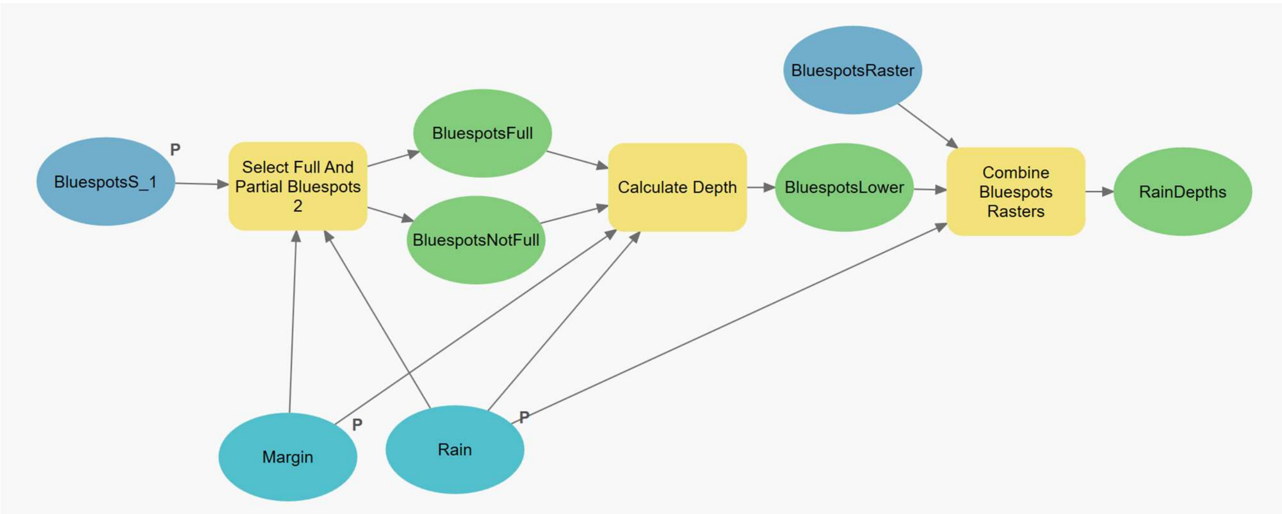


Figure 6-12: Bluespots Depths Assembled 2 model

   a. The *Select Full And Partial Bluespots 2* tool splits the dataset in two portions: Bluespots that are full and the ones that are not. It does that by multiplying the *Watershed Area* (m^2) times *Rain* (mm, converted to m) values. If this volume is larger than the volume of the bluespot, it's considered full. For the full ones we already have a depth raster and don't need any more processing at this point. The partially full ones are saved for the next step.

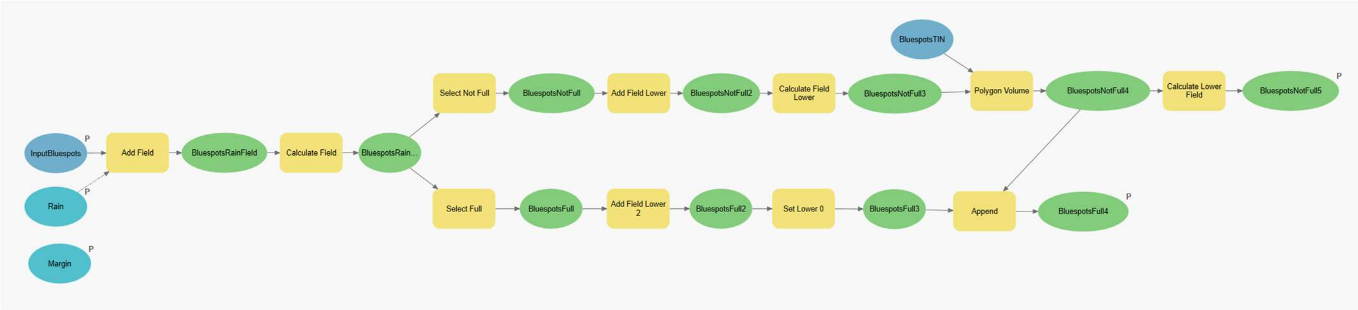


Figure 6-13: Select Full And Partial Bluespots 2

b. The *Calculate Depth* tool calculates the depth (D) of each point in a partially filled bluespot. In this process, it uses the [Polygon Volume](#) tool to calculate the volume between a horizontal plane and the TIN layer with the bluespot depth profile. The tool works iteratively, the horizontal plane gets lowered each step an amount called L based on the calculated volume (V_t) below the plane, the rain volume (V_r) and area of the plane (A_t). Because the Digital Elevation Model always has only one Z-value for each XY-point, the area only decreases when lowering the plane. Thus it can be certain that the plane must be lowered by at least $(V_t - V_r) / A_t$, which is the calculated volume minus the rainfall volume, divided by the calculated area.
Then it calculates the volume and area again (now both smaller) based on the new value of L, and that continues until the calculated volume is smaller than the rain volume times a certain margin (by default 1.01). This way it basically approaches the depth (D) like a limit function. Each time a feature is lowered below the margin, it is removed from the *NotFull* feature and added to the *Full* one.

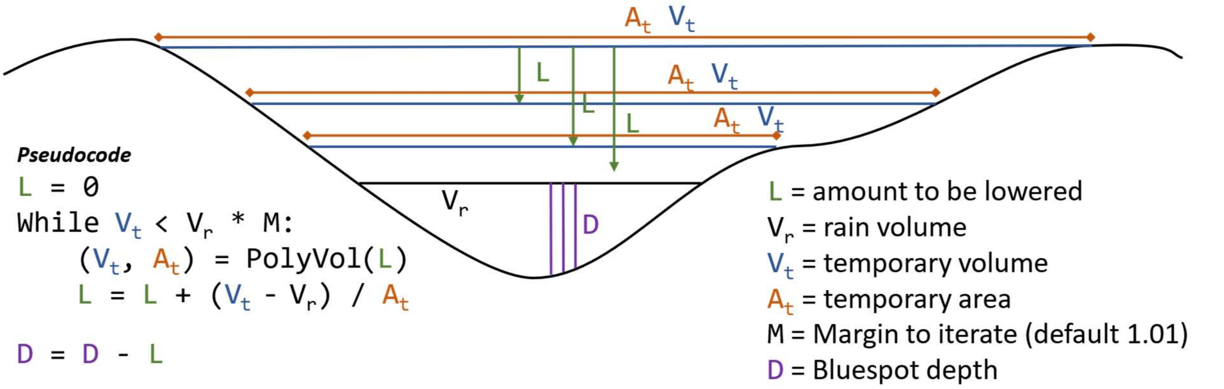


Figure 6-14: Calculate depth illustration (illustration by author)

In each iteration, the *Calculate Depth* tool checks for each feature if the polygon has been lowered enough to fit in the volume margin. If so, the bluespot is moved to the *Full* table. If the table with *NotFull* bluespots it empty and has 0 rows, then the iteration is finished, all bluespots are now lowered enough to fall within margin.

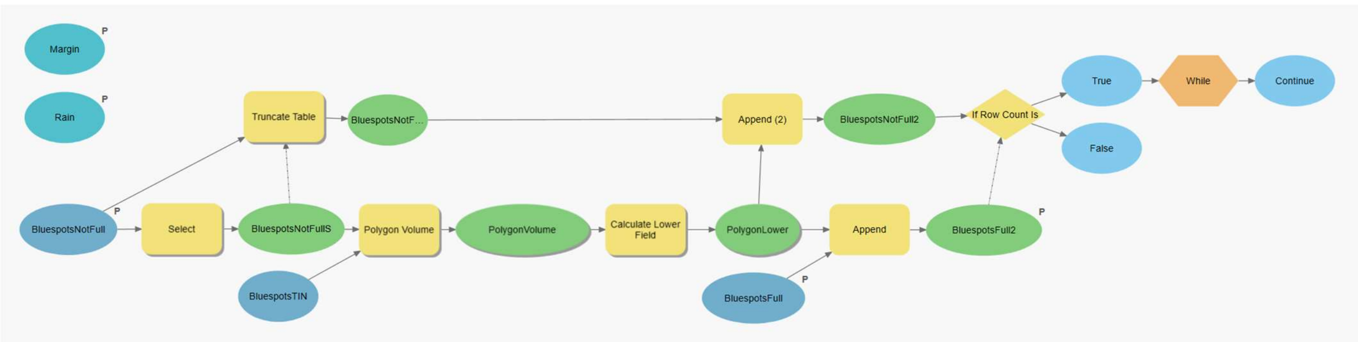


Figure 6-15: Calculate Depth tool

c. At last, the *Combine Bluespots Rasters* tool lowers the bluespot depth raster based on the *Lower* value (called L in figure 8-14) of each polygon calculated in the previous step. All the negative values are set to *NoData*.

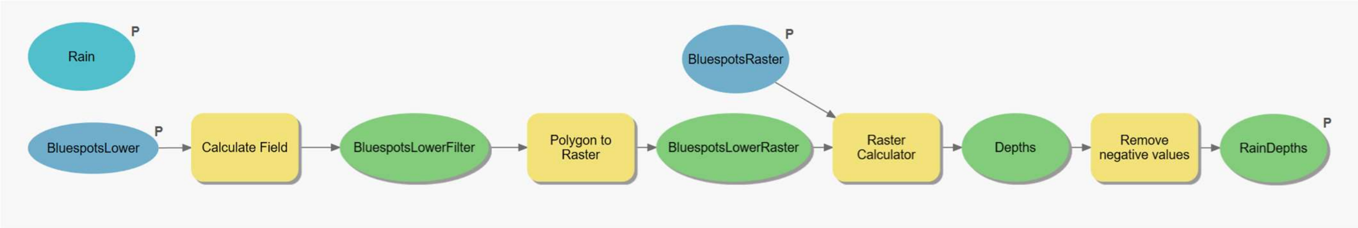


Figure 6-16: Combine Bluespots Rasters tool

3. Finally, the *Add RainDepth Stats to Roads* tool adds information about the bluespot depth to all the intersecting roads. By default, it adds the mean depth, maximum depth and (total) length of the intersecting bluespots.

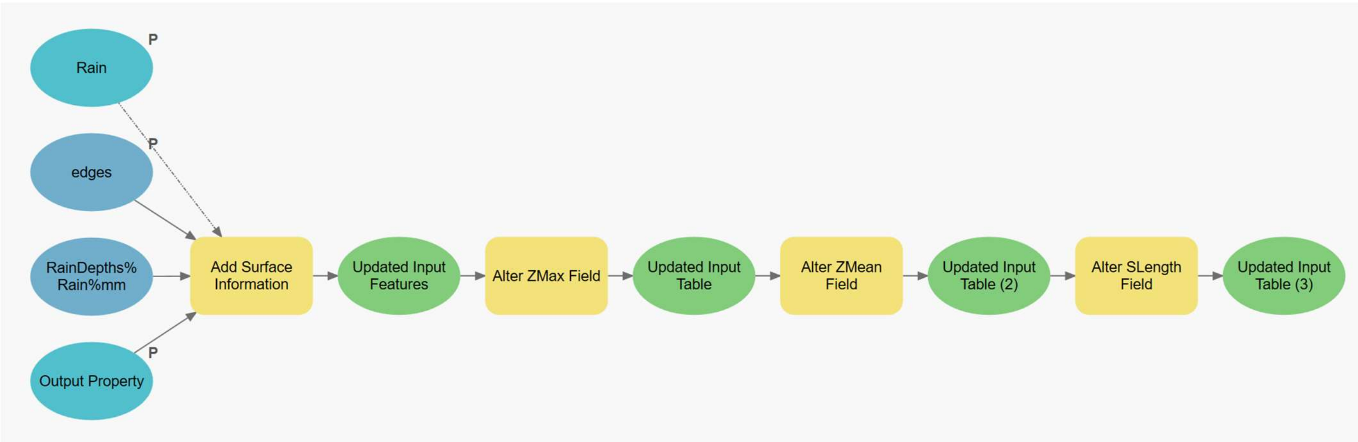


Figure 6-17: Add RainDepth Stats to Roads tool

The result is the *Edges* feature which now has attribute values for each rainfall. These attribute values are saved in a table, which is exported as a CSV file so that it can be imported into the Python notebooks.

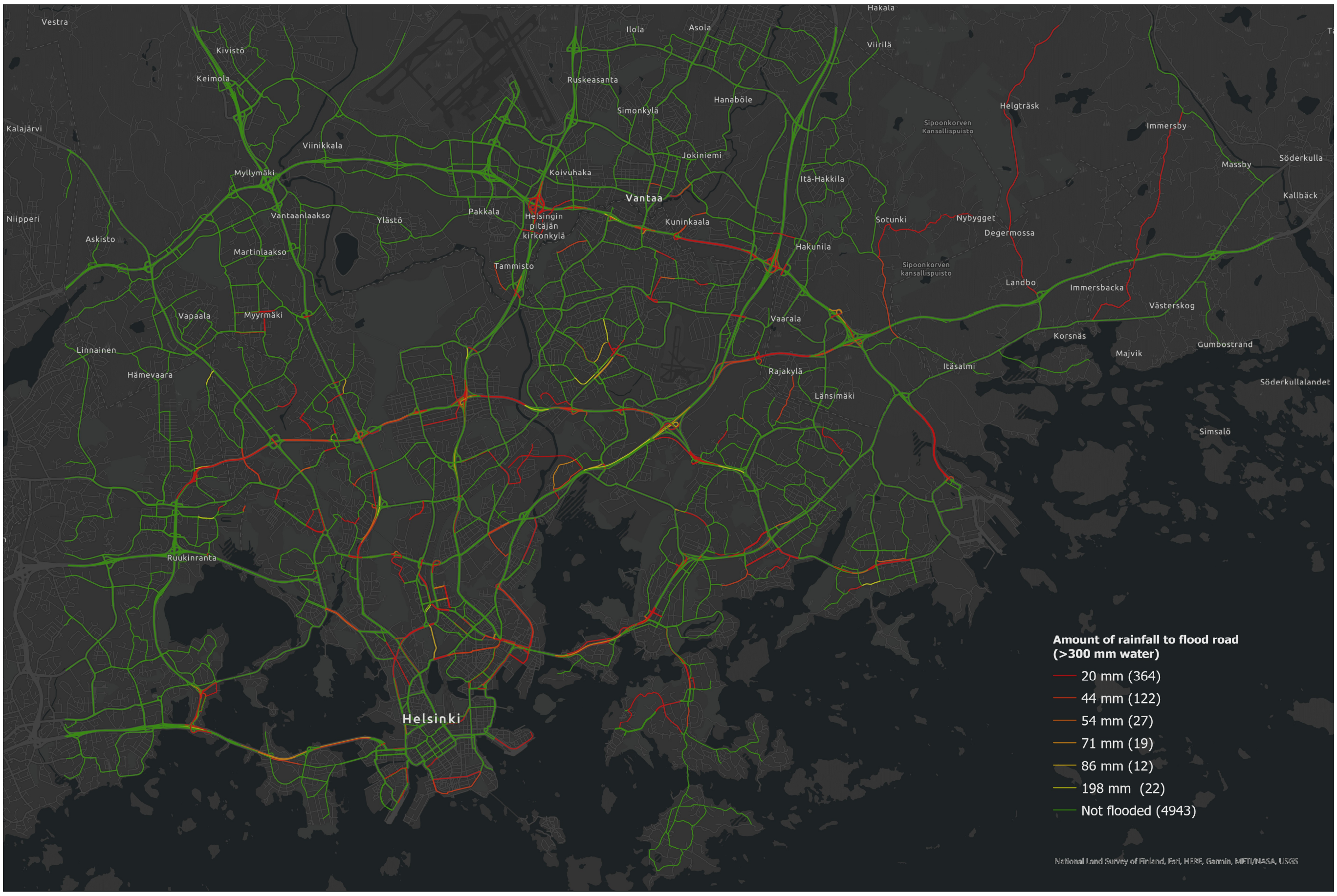


Figure 6-18: Map indicating at which rainfall scenario the road is fully flooded, with a maximum depth of more than 300 mm

The toolbox with all the models used and a minimum ArcGIS Pro package are available here:

- *ArcGIS toolbox: https://github.com/EwoutH/Helsinki-flood-BEP-project/blob/main/stormwater-depth/RainDepthProduction.tbx*
- *ArcGIS Pro package: https://www.arcgis.com/home/item.html?id=5dd6377dc9384185858e078faabbc99d*

Appendix C: Road graph

The road network is the second component, aside from the flood model, that needs to be constructed.

Structure and tools

While multiple structures were possible, as described in Speičys & Jensen (2008), a graph structure was chosen after consideration of the requirements, availability, and scope. The road network should be able to accurately model travel times, while simultaneously have a geospatial component to link a flood risk to each road. There should also be information about speed or travel time for each road.

OSMnx, an open-source Python package developed and maintained by Geoff Boeing (Boeing, 2016) (Boeing, 2017), allows downloading OpenStreetMap data and constructing a NetworkX graph from it. NetworkX is one of the most popular graph analysis packages available. Meanwhile OpenStreetMap (OSM) data is maintained by thousands of open source contributors.

OpenStreetMap contains an up-to-date database of all roads, covers the Helsinki area fully, has attributes for the road class, maximum speed and number of lanes and many more attributes (OpenStreetMap, 2021). Even during this 10-week project, the dataset continually changed with minor fixes, improvements and new road constructed. Thus, the final OSM data for this research was frozen on June 28th 2021, 23:50 UTC+2.

NetworkX meanwhile is a widely used Python package for graph and network analysis. Because Python is a high-level language and NetworkX is written fully in Python, it can be a lot slower than other Python graph packages, such as *graph-tool* or *igraph* (Lin, 2020), that are either partially or fully implemented in a low-level programming language such as C and further optimized. However, for this research it proved fast enough to do all the analysis, which were mostly multi-source shortest path calculations using Dijkstra's algorithm.

The combination of highly accurately and actual data from OpenStreetMap with the possibility to use one of the most used graph packages for Python, NetworkX, made OSMnx in combination with it's dependencies, a great tool to execute this analysis.

Graph construction and pre-processing

The road graph was constructed from OpenStreetMap using OSMnx. The following steps were executed:

- A set of road types to be included was defined. This included all road classes up to *tertiary* roads, as defined in the OSM *highway* key (OpenStreetMap, 2021). Link roads and roads already under construction were also included. Residential, unclassified and other minor road classes were not included to keep the graph to an acceptable size.
- A rectangular bounding box was defined for roads that should be included. It spanned from 60.3285688 to 60.1192574 degrees North and 25.3166706 to 24.7658042 degrees East. The full stormwater area was included, plus a safety margin to make sure alternative routes were possible.
- The graph was projected to the Finland national grid coordinate system, GK25FIN (a.k.a. EPSG:3879).
- The graph was simplified to reduce its complexity. A simpler graph reduces computing needs, both in processing power and time and memory requirements. The effect of this simplification will be further explored in the *Reducing graph complexity* section of this appendix.
  The simplification was done in two steps:
  - All nodes that had only two streets, and thus were just middle points on a line, were removed and the edges were joined together. The length of both edges got summed

up and transferred to the new edge. This was done until only intersections (with 3 or more streets) or dead ends (with exactly one street) remained.
  - Intersections were consolidated by merging all points withing a specified radius of 20 meters into a single node. This way complex intersections such as roundabouts or intersections with multiple nodes could be represented by a single node, reducing the complexity greatly.
- The free flow speed was recalculated by dividing the new length (in meters) by the new travel time (in seconds), and converting the unit from m/s to km/h. This could not be done by default at the time of writing, because the travel time attributes were stored as a set, so an upstream patch to OSMnx was provided and merged (ter Hoeven, 2021b).
- Betweenness centrality was calculated for the graph and added as an attribute to each edge. This provided a baseline for detecting critical roads without the trip weights.

The final graph was saved as GeoPackage and GraphML files, the former to be used with GIS software and the latter to be imported with OSMnx again. The graph contained 2587 nodes, 5580 edges and covered a total of 1439 kilometres of streets, of which some bi-directional, for a total of 1919 km.

All code used in this process can be found in the "OSMnx_road_graph_compilation_BEP.ipynb" Jupyter notebook, available on GitHub (ter Hoeven, 2021c).

	Nodes	Edges	Total edge length (m)	Average edge length (m)	Street segment count	Total street length (m)	Streets per node	Inter-sections
Unsimplified	61734	85221	1915651.3	22.48	64725	1449525.9	2.098	61648
Merged 2-street nodes	5109	9374	1915616.3	204.35	8097	1449491.0	3.182	5023
Consolidated intersections	2587	5580	1919452.1	343.99	4525	1439621.4	1.967	1684

Table 6-2: Graph statistics before and after simplification

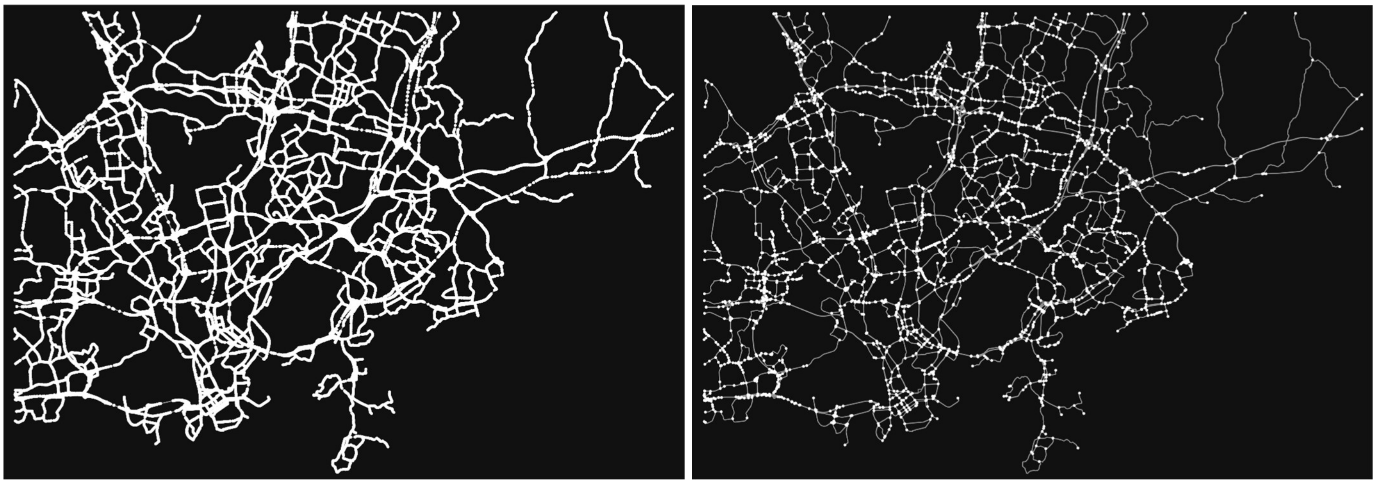


Figure 6-19: The unsimplified graph (left) and fully simplified graph (right)

Reducing graph complexity

From Table 8-2 it can be observed that the number of nodes and edges is greatly reduced after the two simplification steps. This is essential to perform analysis within the computing resources. The execution time of Dijkstra's algorithm, which is used to determine the shortest path between two nodes, scales linearithmicly (a.k.a. N log N) to its number of vertices (nodes) and edges. In Big O-notation the worst-case for a sparse, not strongly connected graph is a time complexity of $O((V + E)\log_2(V))$.

	Vertices (nodes)	Edges	Dijkstra's complexity	Normalized complexity	Speedup
Unsimplified	61734	85221	1621000.4	1.0000	1.000
Merged 2-street nodes	5109	9374	123666.8	0.0763	13.108
Consolidated intersections	2587	5580	64178.4	0.0396	25.258

Table 6-3: Comparison of Dijkstra's algorithm complexity of unsimplified and simplified graphs

The results of calculating this worst-case complexity of Dijkstra's algorithm are shown in Table 8-3. They show the first simplification step reduces this complexity by over a factor 13, and the second step almost halves the complexity again. The combined theoretical speedup is around a factor 25.

Origin-destination data

At this point a graph representing the road network is constructed, from which information how much time it takes to travel from one point (node) to another can be derived using shortest-path algorithms like Dijkstra's. One could simply calculate each possible route and travel time between each pair of two points and average them, but that would not give a representative average of the travel time. What we need is a weighted average, with the weight being the number of vehicles that travel between each set of two points.

Those weights are typically found in origin-destination (OD) data. OD-detail at this level is difficult to get. Congestion levels are available at the edge/node level, such as *opentraffic* (opentraffic, 2021) and Esri's *World Traffic Service* (Esri, 2021), those only contain information about traffic speeds and not traffic volumes.

OD data was found available from TomTom, with a few limitations:

- A maximum of 4 date ranges can be defined
- At least 7 days should be in each date range
- A maximum of 4 time ranges can be defined
- Total size of regions is limited to 3000 km²
- A maximum of 200 regions can be defined
- The data covers January 2020

Since the stormwater raster only covers 255 km^2, and the road network bounding box is just slightly bigger, the 3000 km^2 region area limit did not pose a problem. The maximum number of 200 regions was a limitation though, since each of the 2587 nodes could get its own OD data.

The approach chosen was to group nearby nodes together to form a smaller number of regions. The *Generate Subset Polygons* tool was used in ArcGIS Pro to create regions that each contained a minimum and maximum number of nodes. On average, each region should contain 12.935 nodes. Experimentation proved that a minimum of 12 and maximum of 16 nodes per region resulted in 199 regions, below the 200 TomTom limit. A map with the defined OD regions is presented in Figure 6-20.

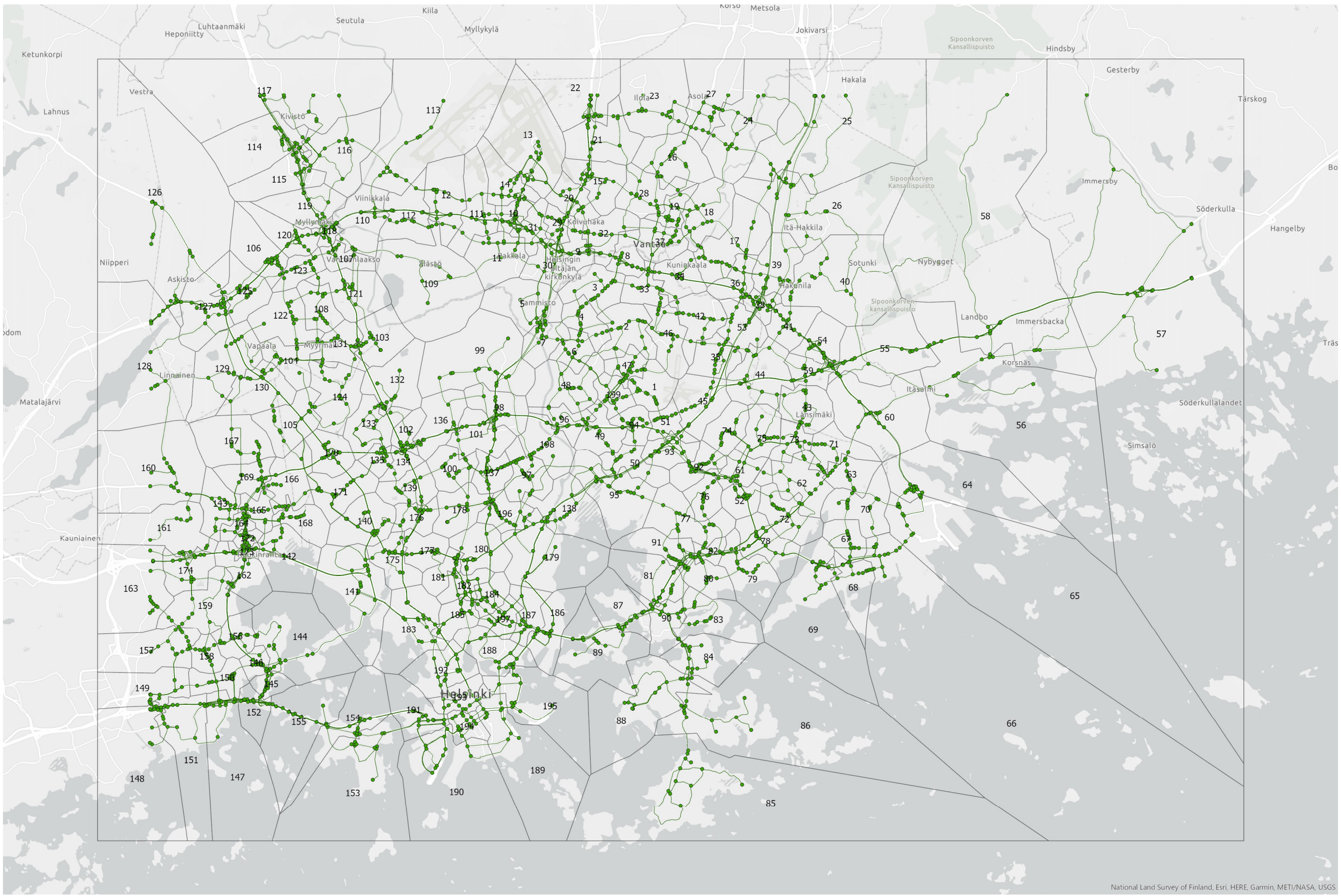


Figure 6-20: The 199 regions with each 12 to 16 nodes

These polygon regions were subsequently imported in TomTom Move. Two datasets were generated, one included all trips that crossed through one or more of the regions, representing all traffic, and one that only included traffic that both started and ended in one of the defined regions. This way both the full impact on travel time can be measured, and just the impact on trips within Helsinki.

The date range chosen was January 4 to 31, 2020. This way exactly 4 weeks were included, the maximum number of whole weeks without holidays to fit in the date range. Each weekday occurs the same number of times.

Spatial Sankey diagrams are displayed in Figure 6-21 (all traffic) and 6-22 (internal traffic), with darker and more red lines representing more traffic.

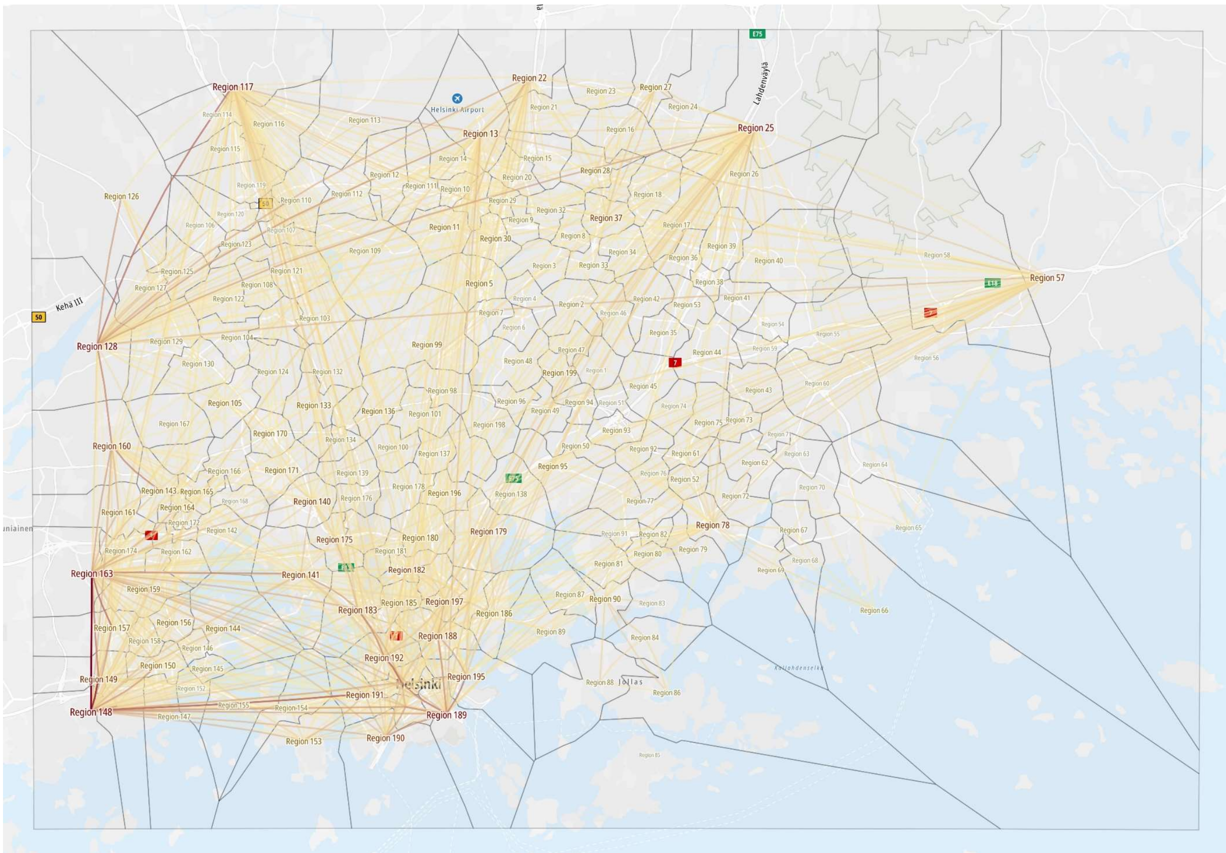


Figure 6-21: Spatial sankey diagram of traffic volume for all trips through the Helsinki region (TomTom, 2021a)

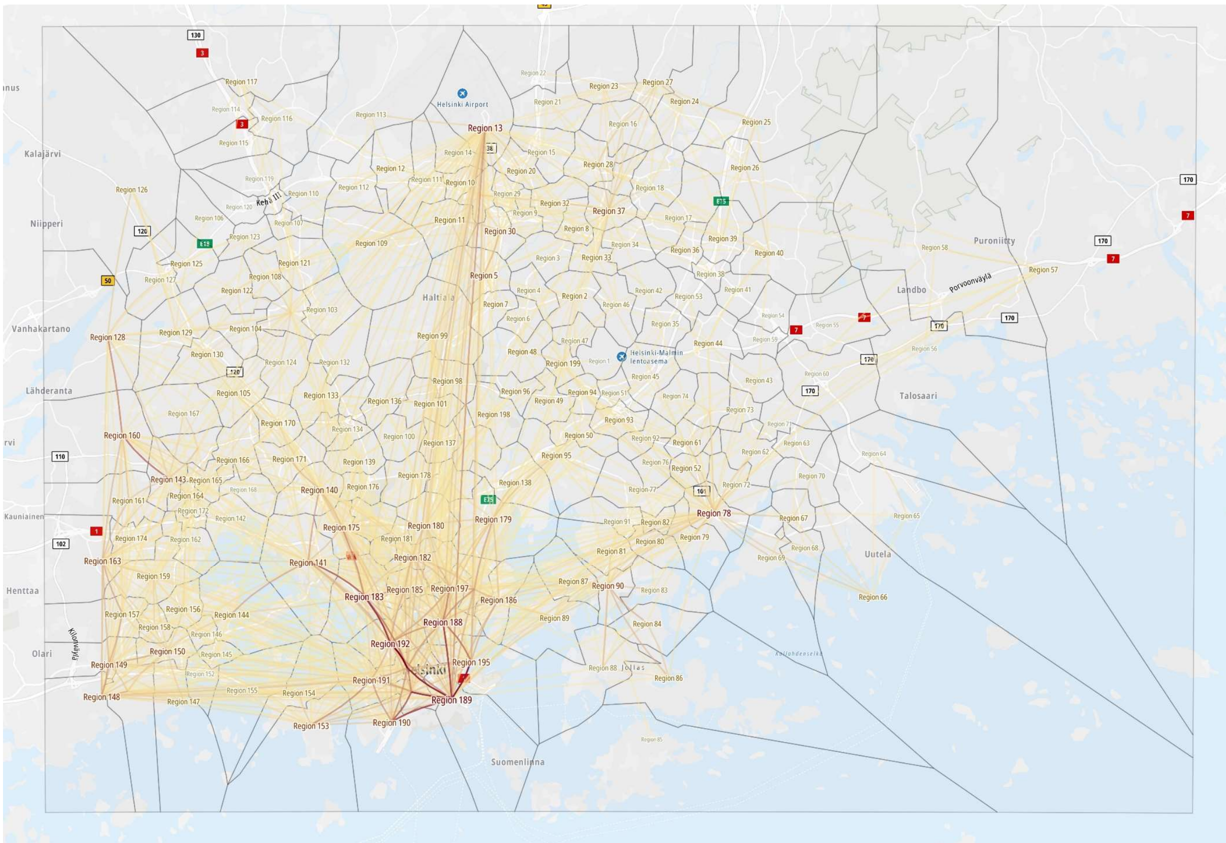


Figure 6-22: Spatial sankey diagram of traffic volume for internal trips in Helsinki (starting and ending in the defined region) (TomTom, 2021b)

This OD data was exported to CSV files to be loaded into the Python notebooks.

Driving speed

Some roads where not flooded over the limit of what's considered drivable but were flooded to some extent. On these roads the regular maximum speed was not reachable, so this speed and the derivative travel time needs to be adjusted.

Pregnolato et al. (2017) estimated a depth-disruption function, that links the water depth on a road to the driving speed. The formula is as follows:

$$v(w) = 0.0009w^2 - 0.5529w + 86.9448$$

in which *w* is the water depth in millimeters and *v(w)* the vehicle speed in km/h. According to Pregnolato et al. (2017), the fitted function has an R-squared value of 0.95. The function is valid in the range of 0 to 300 millimetres water depth. The speeds for this range are plotted in Figure 8-23.

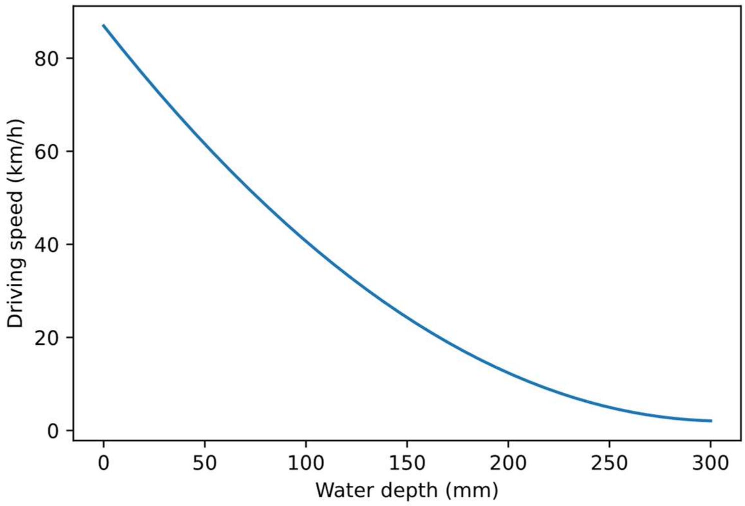


Figure 6-23: The relation between water depth and driving speed according to Pregnolato et al. (2017)

Most road segments are that are flooded are not flooded over the full length, but only on certain parts. Since on the other parts the regular maximum speed can be driven, our calculation take that into account. The maximum depth is used to determine if the road is flooded or not at all, the mean depth and length of the portion flooded are used to determine the new travel time. For the flooded length of the road the depth-disruption function is used to calculate maximum speed and from there the travel time, while for the length of the road that isn't flooding the regular maximum speed is used to calculate the travel time for that portion. Both travel times are then summed together.

This process results in the following formula for the new average driving speed.

$$v_a = \max(v_f, v_r) * \frac{L_f}{L_r} + v_r * (1 - \frac{L_f}{L_r})$$

In which:

L_r is the total length of the road (in m)
v_r is the regular driving speed on the road (in km/h)
L_f is the length of the portion of road flooded (in m)
v_f is the speed on the portion of road flooded (in km/h)
v_a is the resulting average speed (in km/h)

Then the road segment length is divided by this new average speed, to get the travel time of the partially flooded road segment.

Appendix D: Extended results

This appendix contains additional and extended result tables and figures to support the results section.

Most used road segments

The following table is a more extensive version of Table 4-1 and lists the most used road segments in Helsinki. Note the Kehä I highway and Tuusulanväylä motorway containing the most used road segments.

Road segment	**Name**	**Type**	**Weighted centrality (ext)**	**Weighted centrality (int)**	**Used in % of trips (ext)**	**Used in % of trips (int)**
(1459, 1464)	Kehä I	trunk	77572.3	50291.0	5.545%	3.595%
(1794, 131)	Kehä I	trunk	72298.2	49862.6	5.168%	3.564%
(589, 2251)	Kehä I	trunk	72268.3	48079.6	5.166%	3.437%
(21, 20)	Tuusulanväylä	motorway	72065.4	59889.6	5.152%	4.281%
(549, 550)	Länsiväylä	motorway	71457.2	21578.0	5.108%	1.543%
(280, 505)	Lahdenväylä	motorway	71029.7	36084.1	5.078%	2.579%
(210, 207)	Kehä I	trunk	70612.7	46069.5	5.048%	3.293%
(1465, 210)	n/a	trunk_link	70527.9	45995.3	5.042%	3.288%
(206, 1465)	n/a	trunk_link	69871.5	45551.9	4.995%	3.256%
(218, 206)	Kehä I	trunk	69786.7	45477.7	4.989%	3.251%
(129, 1459)	Kehä I	trunk	68509.6	42122.1	4.897%	3.011%
(1537, 1438)	Tuusulanväylä	motorway	67757.2	55490.7	4.844%	3.967%
(841, 21)	Tuusulanväylä	motorway	65097.7	53655.5	4.654%	3.836%
(793, 1635)	Tuusulanväylä	motorway	61984.0	48990.9	4.431%	3.502%
(20, 341)	Kehä I	trunk	59000.1	47858.0	4.218%	3.421%

Table 6-4: Most uses road segments

The full dataset is available on the GitHub repository in the *proc_weighted_centrality.xlsx* spreadsheet.

Unweighted travel time

The following figures show the travel time for each route for different scenarios. In the main report only the weighted averages are discussed, but full data is available for other statistics to be calculated.

A histogram (Figure 6-24) and a boxplot (Figure 6-25) are displayed below. For the boxplot, the orange line is the median and the green triangle the average.

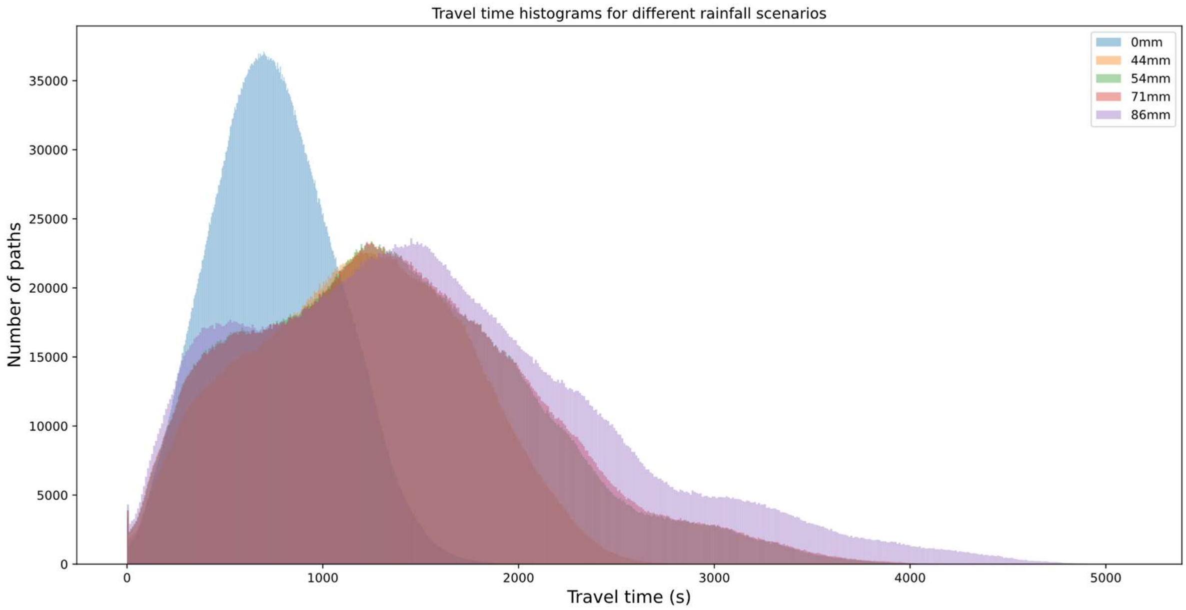


Figure 6-24

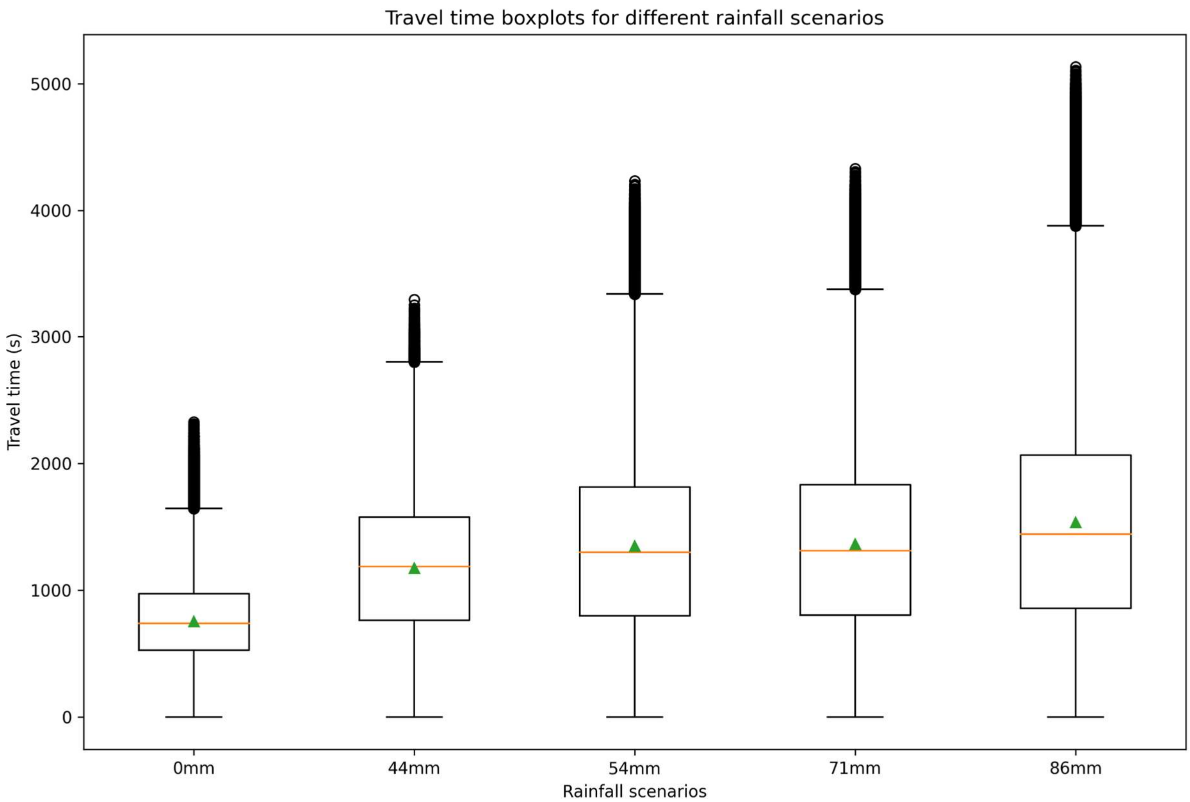


Figure 6-25

Weighted travel time

All traffic

The following two sets of three figures show the travel time weighted with the number of trips. Figure

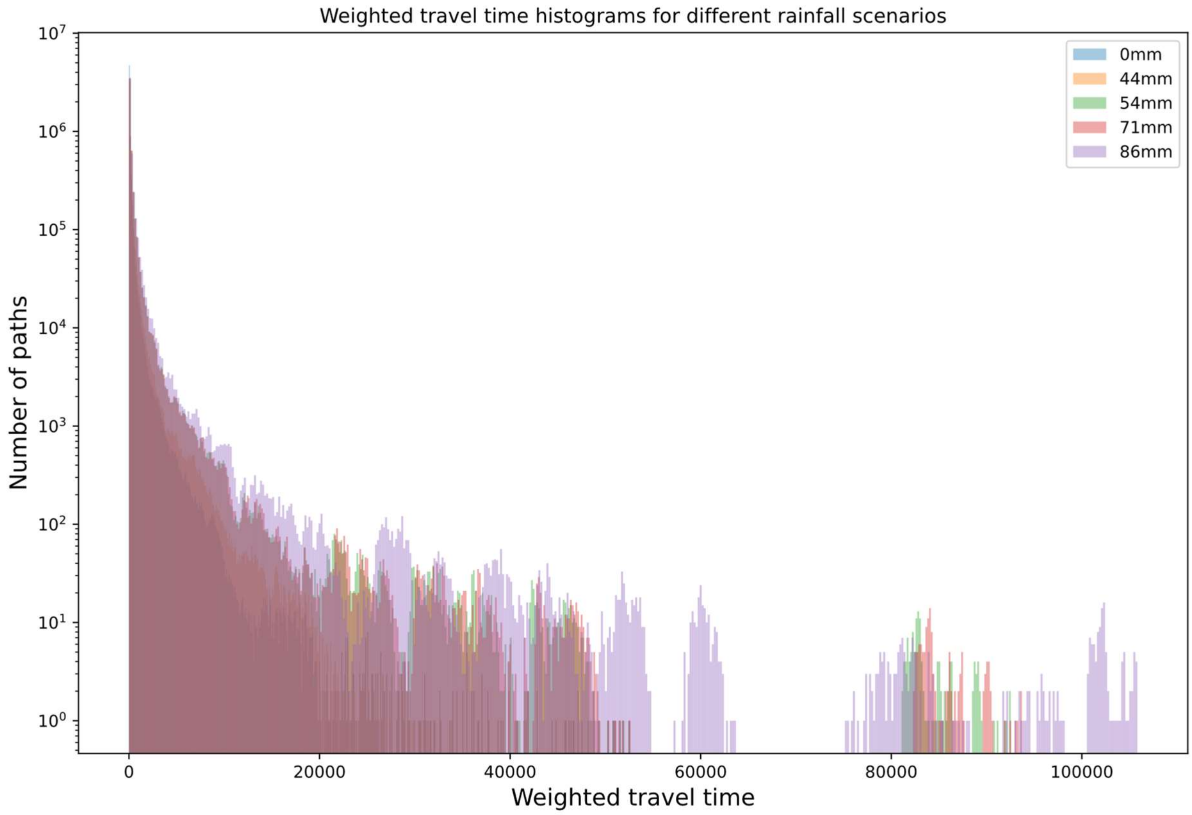


Figure 6-26

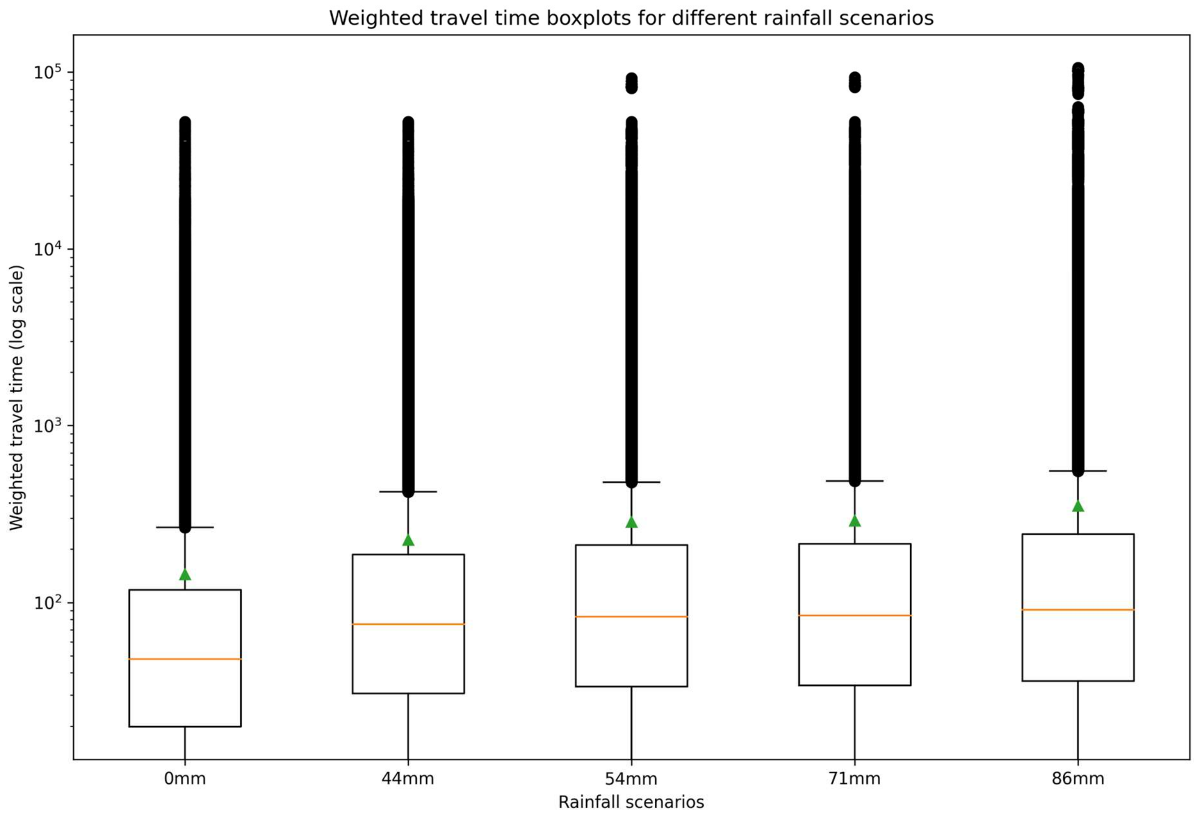


Figure 6-27

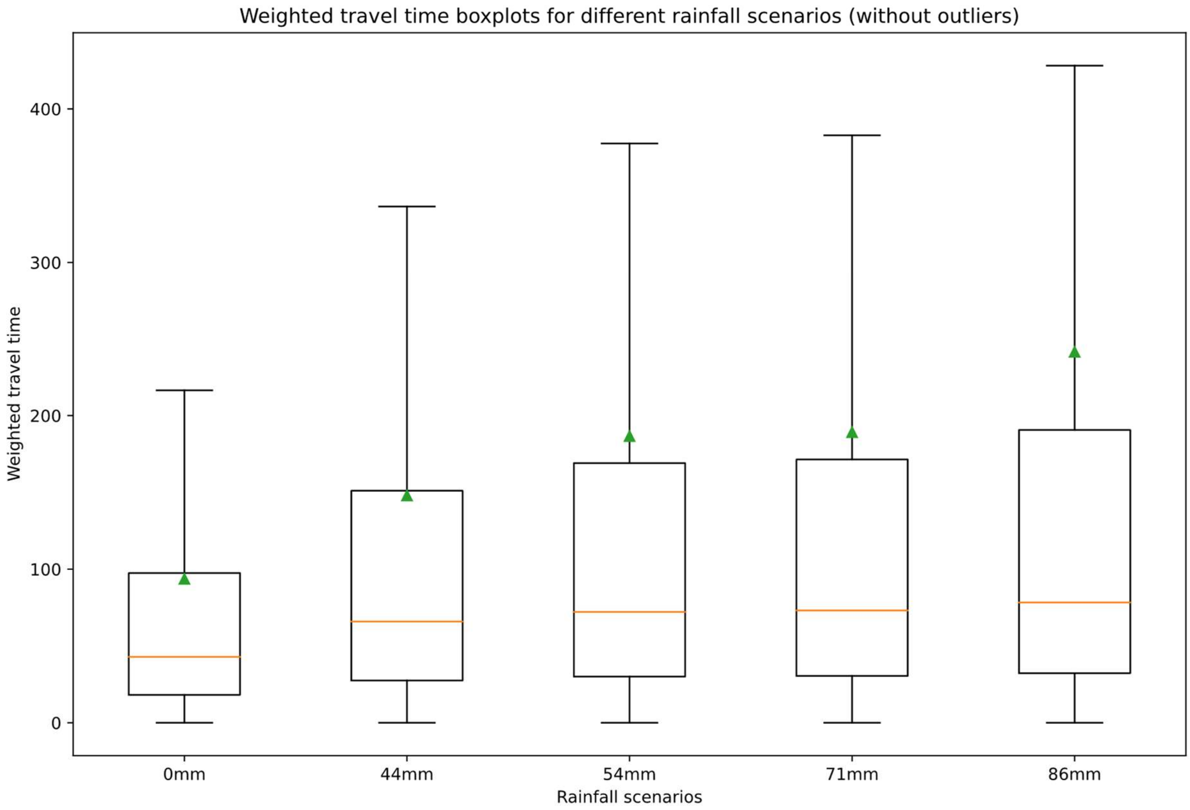


Figure 6-28

Internal traffic only

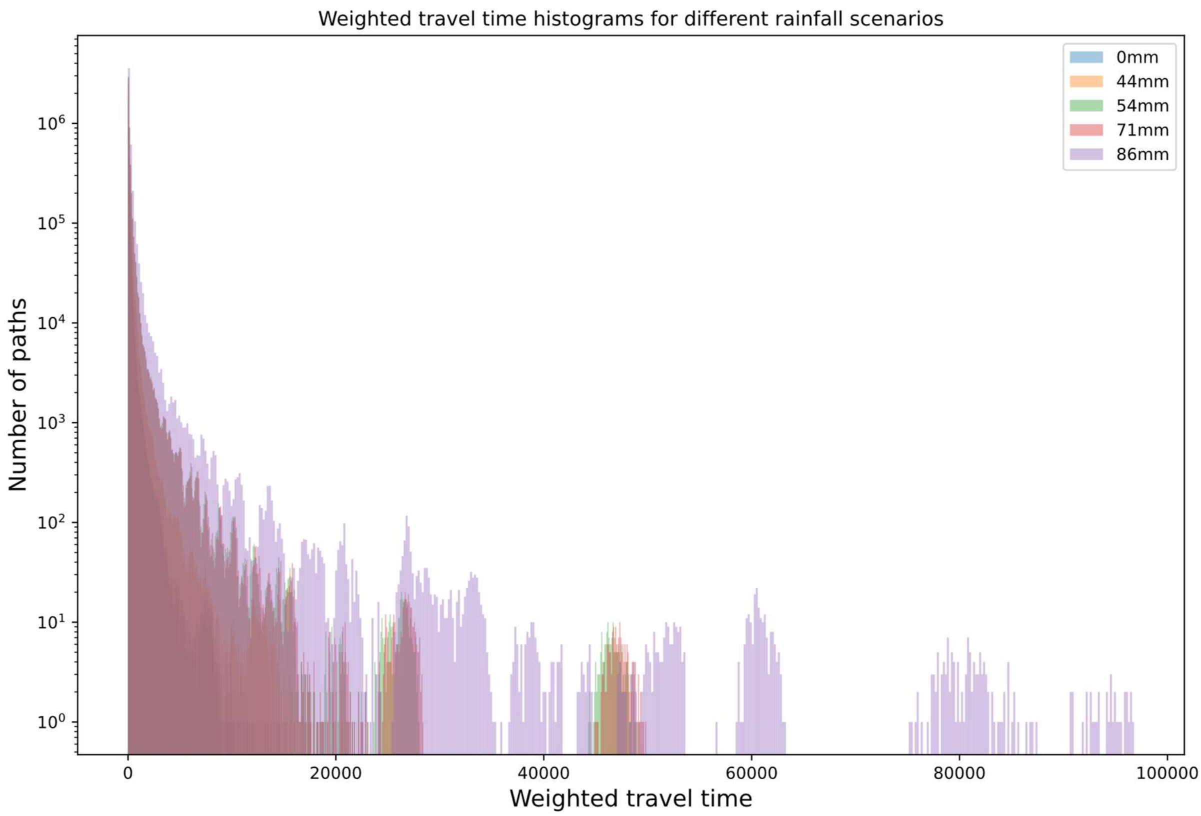


Figure 6-29

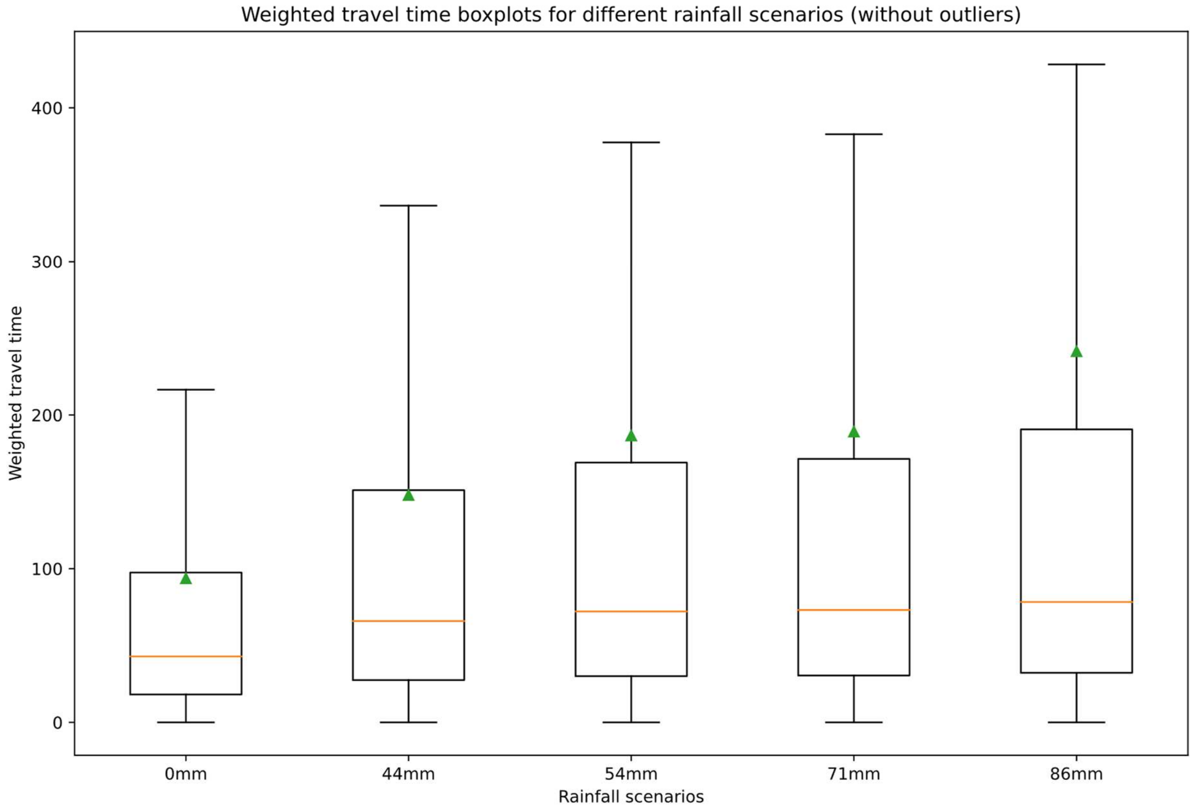


Figure 6-30

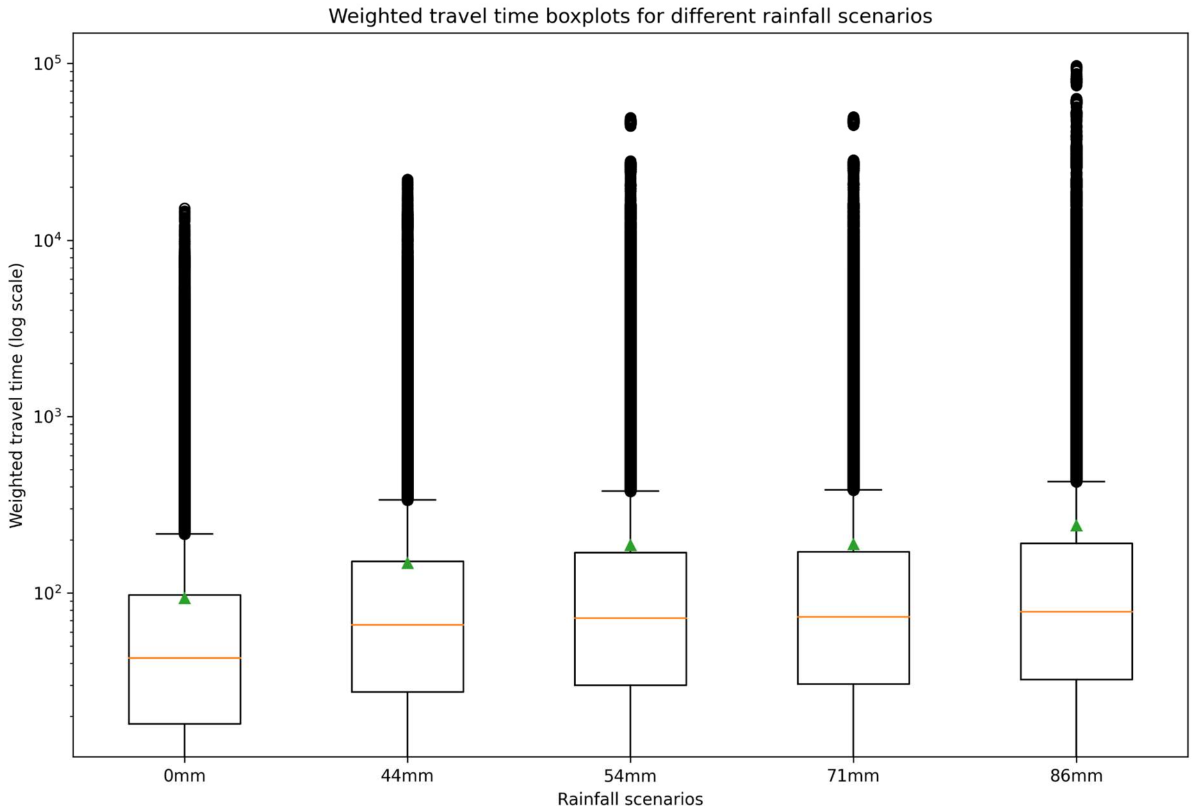


Figure 6-31

Travel time increases of single flooded road segments

This table contains the full travel time delays when single road segments are removed from the graph. The first two color-coded columns states the delay for the external traffic dataset, the last two for the internal traffic only.

Only the 12 highest risk road segments for each rainfall scenario for each dataset were measured for each dataset, so some road segments have missing values due to falling within the 12 highest risk for one dataset, but not the other.

The *Flooded at* column contains the rainfall scenario at which a road starts flooding to an amount that is becomes undrivable (>0.30 m). For example, a road segment that floods at 44 mm, also floods in the higher rainfall scenarios.

Road segment	Name	Type	Flooded at [mm]	Ext. Delay [s]	Ext. Increase	Int. Delay [s]	Int. Increase
(1762, 1851)	['Kehä I', 'Mestarintunneli']	trunk	20	2.846	0.579%		
(2030, 220)	Kehä I	trunk	20			2.373	0.545%
(324, 1686)	Sörnäisten rantatie	primary	20			1.842	0.423%
(1463, 130)	Kehä I	trunk	44	5.011	1.019%	4.819	1.106%
(214, 1409)	Kehä I	trunk	44	4.101	0.834%	3.548	0.814%
(18, 1977)	Kehä I	trunk	44	3.526	0.718%		
(540, 17)	Kehä I	trunk	44	3.502	0.712%	2.811	0.645%
(281, 280)	Lahdenväylä	motorway	44	2.818	0.573%		
(1413, 126)	Kehä I	trunk	44	2.234	0.454%	1.859	0.427%
(38, 43)	Kehä III	trunk	44	1.028	0.209%	0.846	0.194%
(43, 42)	Kehä III	trunk	44	0.860	0.175%	0.640	0.147%
(1635, 12)	Tuusulanväylä	motorway	44	0.775	0.158%	0.527	0.121%
(2251, 218)	Kehä I	trunk	44	0.700	0.142%	0.607	0.139%
(130, 1794)	Kehä I	trunk	44	0.173	0.035%	0.156	0.036%
(1448, 588)		motorway_link	44	0.161	0.033%	0.199	0.046%
(516, 814)	Hämeenlinnanväylä	trunk	44	0.002	0.000%	0.001	0.000%
(150, 153)	Länsiväylä	motorway	54	8.215	1.672%	5.821	1.337%
(218, 206)	Kehä I	trunk	54	4.844	0.985%	4.459	1.023%
(127, 128)	Kehä I	trunk	54	3.634	0.739%	3.347	0.768%
(1039, 1038)		trunk_link	54	0.346	0.070%	0.373	0.086%
(1777, 314)		primary_link	54	0.087	0.018%	0.088	0.020%
(149, 150)	Länsiväylä	motorway	71	1.416	0.288%		
(146, 2515)	Länsiväylä	motorway	71	0.996	0.203%	0.735	0.169%
(918, 919)	Kotinummentie	secondary	71	0.026	0.005%	0.033	0.008%
(2194, 2514)	Länsiväylä	motorway	86	1.994	0.406%	1.469	0.337%
(1409, 540)	Kehä I	trunk	86	0.873	0.178%	0.665	0.153%
(381, 331)	Pihlajamäentie	secondary	86	0.365	0.074%	0.466	0.107%
(331, 1688)		motorway_link	86	0.346	0.070%	0.460	0.106%
(956, 957)	Tapanilankaari	secondary	86	0.182	0.037%	0.228	0.052%
(957, 1256)	Tapulikaupungintie	secondary	86	0.140	0.029%	0.183	0.042%
(941, 918)	Malmin raitti	tertiary	86	0.032	0.006%	0.040	0.009%
(1638, 1342)		motorway_link	86	0.023	0.005%	0.029	0.007%

(589, 2251)	Kehä I	trunk	198	4.429	0.901%	4.029	0.925%
(1687, 35)	Lahdenväylä	motorway	198	2.056	0.418%	1.445	0.332%
(1742, 841)		motorway_link	198	1.904	0.387%	1.904	0.437%
(2201, 216)	Vihdintie	primary	198	1.624	0.330%	1.501	0.345%
(208, 2201)	Vihdintie	primary	198	1.495	0.304%	1.418	0.325%
(35, 36)	Lahdenväylä	motorway	198	1.374	0.279%		
(209, 2201)	Vihdintie	primary	198	1.368	0.278%	1.258	0.289%
(2201, 209)	Vihdintie	primary	198	1.230	0.250%	1.166	0.268%
(1402, 1401)	Vaskivuorentie	secondary	198	0.603	0.123%	0.747	0.172%
(49, 1040)		trunk_link	198	0.371	0.076%	0.411	0.094%
(1827, 1794)		trunk_link	198	0.280	0.057%	0.346	0.079%
(515, 1402)		trunk_link	198	0.269	0.055%	0.277	0.064%
(1763, 233)	Turuntie	secondary	198	0.263	0.053%	0.287	0.066%
(2514, 149)	Länsiväylä	motorway	198	0.178	0.036%		
(2392, 1717)	Lemissaarentie	primary	198	0.140	0.028%	0.174	0.040%
(576, 1222)		trunk_link	198	0.136	0.028%	0.179	0.041%
(1717, 2530)	Lauttasaarentie	secondary	198	0.097	0.020%	0.097	0.022%
(165, 160)	Länsiväylä	motorway	198	0.086	0.017%	0.060	0.014%
(1543, 1777)		primary_link	198	0.027	0.005%	0.032	0.007%
(919, 920)	Kotinummentie	secondary	198	0.018	0.004%	0.023	0.005%
(1636, 159)	Länsiväylä	motorway	198	0.014	0.003%		

Table 6-5: Average travel time increases caused by single flooded road segments

At the 44 mm scenario, road segments of the Kehä I (Ring 1), the inner-most highway ring of the Greater Helsinki region, are flooded. They lead to large increases in travel time, with the (1463, 130) segment, just West of Länsi-Pakila, increasing average travel times with 1.019% and 1.106% for the full and internal only traffic dataset respectively.

With 54 mm of rain the (150, 153)-segment on the Länsiväylä (Western Highway) starts to flood. This motorway connects the center district with the western parts of Helsinki, including the Lauttasaari island.