Data quality of the Bangladesh infrastructure database

Group 13:
Thomas Sandbergen 4720814
Thomas Rous 4963946
Matthijs van de Wiel 4896947
Yassine Mouhdad 5879566
Tom Hillenaar 4861973

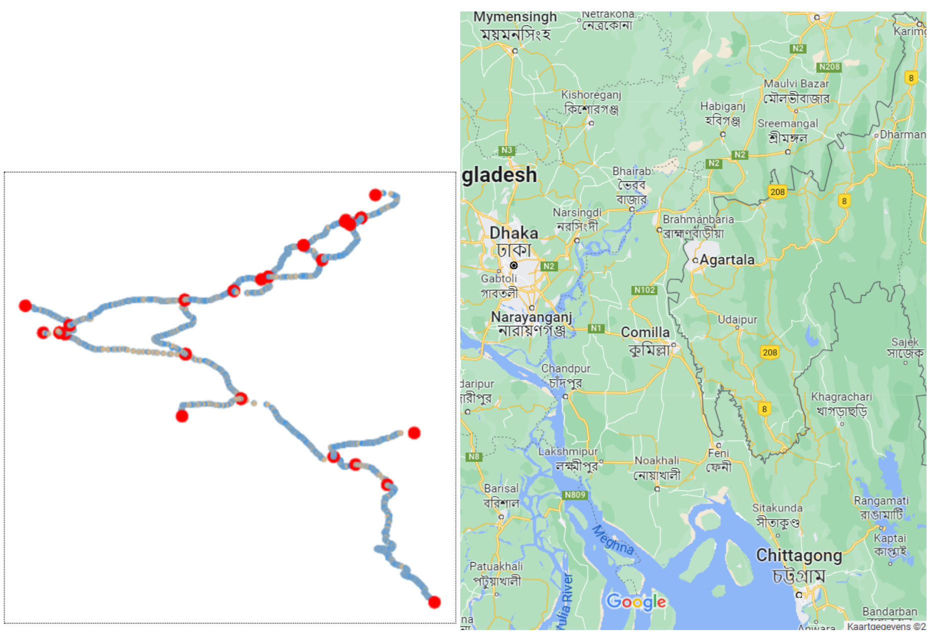
April 7, 2023

1 Introduction

The goal of the task given by the World Bank is to find out critical and vulnerable parts of the infrastructure of Bangladesh that should be invested in. Natural disasters can cause critical infrastructure to get damaged and break which has a great impact on the country's logistics and the general economy. The road infrastructure in Bangladesh is a massive connected network of roads, bridges, and culverts which asks for a complex model. In order to accurately analyze this system, vehicles are simulated driving the roads to see how much delay time is formed. However, vulnerability and criticality are highly debated in the literature for simulation. Therefore, first, these concepts are analyzed to find out how they can be used in our assignment and which data is needed to be collected to interpret these terms as appropriately as possible for the task given.

There are different possibilities for modeling such complex problems with interaction between agent sets. One of them is agent-based modeling (ABM). The focus in ABM is on individual agents (Klabunde, 2016). This includes their interaction with other agents, the effects of that interaction, and how

this will affect model outcomes. The different interdependencies in transportation (roads, bridges, and vehicles) are complex and ABM allows for micro-data to be micro-simulated (Macal & North, 2005). Since the infrastructure concerns many different instances interacting on a micro-scale that make up the whole infrastructure network, ABM should be appropriate in this case. The 'Mesa' module is used in Python for ABM modeling.
Within this research, the project scope is the N1, N2, and their respective side roads larger than 25 kilometers. Connecting these roads creates an infrastructure network that includes trucks driving to different destinations in both directions. The assignment has been carried out such that it can be generalized to the whole infrastructure network. This assignment highly builds on the previous assignments where the data has been cleaned and simulation has been done with the mesa module.
In order to give a better insight into the road system integrated into the model, the figures below are shown. The left figure shows the output of the visualization part of the model and the right figure shows how these roads are located in the actual world via a screenshot of Google Maps.



((a)) Overview road system ((b)) Road system in Google Maps

Figure 1: Road system of Bangladesh integrated in the model

First, In Chapter 2 criticality and vulnerability are reviewed and how it is used in this assignment is elaborated on. In order to get the most useful representation of the road infrastructure network, the available data has been cleaned and prepared. However, the data is not yet fit for the simulation model as it is. In Chapter 3, the data preparation procedure is explained to fit the data according to the requirements. Chapter 4 gives elaboration on the model structure needed for this assignment. Chapter 6 explains the outcomes of the vulnerability of the bridges and critical points in the network. Chapter 7 contains the discussion within the report and includes an explanation of assumptions made, potential improvements, and limitations of the model.

2 Criticality & Vulnerability

Criticality and vulnerability are terms often used in the road network and transport analysis field within different contexts. All these contexts apply different definitions of these terms depending on how they want to use them.

In Erath, Birdsall, Axhausen, and Hajdin (2009), the Swiss road network is assessed with a vulnerability analysis that uses the following definition instead of the standard one in transport literature: "a combination of the probability of occurrence of a given hazard, the resistance of the infrastructure against it, and its consequences to transport". The usual definition in transport literature describes vulnerability as increased travel length and decreased accessibility according to Erath et al. (2009). Berdica (2002) refers to the vulnerability of the transport system as problems of reduced accessibility that create disturbances in traffic. Vulnerability is regarded as the most adverse impact on network performance measured in network travel time or generalized travel costs in Luathep, Sumalee, Ho, and Kurauchi (2011). These are just a few examples of slightly different definitions of vulnerability. These differences matter once they are implemented in the code. Therefore, vulnerability is defined extensively in this section.

As seen in the previous paragraph, in road network analysis, the definition of criticality and vulnerability varies across the literature with a focus on the use of the system rather than the actual network itself (Berdica, 2002). In this case, the use of the road network is carrying goods vital to the economy and society across Bangladesh with trucks. This gives a more logistic viewpoint. With this use of the road network in mind, the definition and usage of the terms criticality and vulnerability in this report will be elaborated on in this section.

Following the research of Jenelius and Mattsson (2015), this vulnerability analysis is inherently linked to risk analysis. They follow the line of reasoning from Kaplan and Garrick (1981) that risk can be presented as a list of "triplets": a description of a particular scenario, the probability of that scenario occurring, and the impact of the scenario. In this case, the particular scenario is a bridge breaking down. The probability of this scenario occurring is found in the condition of each bridge and the percentage of bridges for each category of conditions in the experimental setup. At last, the impact of the bridges breaking down is vehicles being delayed more while traveling across a certain route. For logistics, losing time while transporting is essentially losing money, as more successful deliveries yield more income. This shows that this definition of vulnerability accurately follows the use of the road network system in this case.

To mitigate the impact of these vulnerable bridges on the road network system, it is essential to identify the critical links to other parts of the system (García-Palomares, Gutiérrez, Martín, & Moya-Gómez, 2018). The aim of criticality analysis in a transport network is to list the infrastructure components based on their role in the performance of the overall infrastructure network (Jafino, Kwakkel, & Verbraeck, 2019). Several criticality metrics are presented across different cases in literature as shown by Table 1 in Jafino et al. (2019). To accurately suit the use of the system in this case, the criticality metric of choice here is 'Change in unweighted total travel cost' and 'Traffic flow'. More specifically, for every bridge, the delay time it has caused for the trucks will be measured and the number of trucks passing the bridges is calculated. This will show which bridges are most critical to the road network system in Bangladesh.

With these definitions of criticality and vulnerability, the results of this report will give an insight into bridges that are most likely to break down under different scenarios and how critical these bridges are to maintain a good road network for the least impact on the logistics industry. The process of providing these insights will follow in the next sections.

3 Data Preparation

Before the model is used, data preparation was necessary in order to collect and reshape data in the correct format for the model. In order to give a structured overview of the data preparation a table is given with a row for each part of the data preparation process. Notice that many steps are the same as in assignments 2 and 3.

However, as explained in the literature section, data on the number of trucks between each section is needed. Separate files contain data on different vehicles driving between different LRP points. The second table explains what has been done to create useful data from these files. The goal of this process in data cleaning was to find out the number of trucks between different intersections and sourcesinks. The basic concept is that a bridge that serves more trucks in a network is much more critical than a bridge that does not see much traffic.

Table 1: Overview of cleaning within data preparation process

Subject	**Method of data cleaning**
Understanding of essential information for roads and bridges in the simulation model	Eight columns are selected to construct datasets useful for simulation, containing content of the road, id, model type, condition, name, lat, lon, gap, and length.
Selecting datasets	The file 'roads3.csv' contains most of the information necessary for the final dataset. However, bridge-type information is missing from this dataset. Information on bridges is retrieved from 'BMMSoverview.xlsx'.
Create correct length for roads.	Take the differences in chainage values and multiply this number by 1000 in order to get the length in meters.
Selecting side of bridge	Each bridge contains two points: Bridge start (BS) and bridge end (BE). We selected only BS and, therefore, dropped BE.
Dropping column 'gap'	The column 'gap' is not necessary anymore and, therefore, is dropped out of the dataframe.
Renaming column	In order to merge the datasets later on, the column lrp in the roads dataframe has been renamed into 'LRPName'.
Drop duplicates	Duplicate bridges are dropped out of the dataframe. The column used is 'LRPName'.

Merging the two datasets and selecting the necessary columns again	The dataset is mainly based on the 'roads3' file since this file is the most extensive file of the two sets selected. The idea is to add specific information on bridges. For merging the two datasets, the columns 'road' and 'LRPName' are indicators. Some columns are duplicated in the merged dataframe, since both dataframes contain these columns. Therefore, some duplicate columns from origine the BMMSoverview file are being dropped. Important columns added from 'BMMSoverview' are: damage conditions and length of bridges. The column length, however, is also present in the roads file. These columns are, therefore, combined into one column since information is necessary from both columns.
Selecting roads	Selecting the roads within the 'road' column. Each road which starts with either N1 of N2 will be selected. After this process. Every road shorter than 25 km is dropped out of the dataframe.
Completing conditions of bridges	Every bridge should contain a condition (either A, B, C or D). However, some bridge do not contain a condition. These bridges will have assigned conditions. In order to do this, the distribution between bridge conditions has been used. Each bridge with missing condition will have a probability to have either condtion A, B, C or D assigned.

Implementing ID number	The dataframe does not contain an id number for each part of the network. This is necessary for the model to run and, therefore, has been added to the model. The first datapoint has id 0, the second id 1 etc.
Renaming type of bridges	There are currently either 'Culvert' or 'Bridge' type of bridges in the dataframe. Both type of bridges are implemented in the model and will have the 'Bridge' type.
Applying the setup of the demo file	The column structure of the demo file has been applied on the dataframe in order to make sure that the model can run with NetworkX
Renaming specific datapoints into either sourcesinks, intersection or link	An important asset in the third assignment is the column: 'type'. However, after scanning the data set, it could be seen that several changes had to be made. For instance, if 'CrossRoad' was indicated it must be changed to 'Intersection'. Another example is the change from 'Culvert' and 'Bridges' to just 'bridges'. In this way, 'type' only consists of the following four values: 'sourcesink', 'link', 'bridge', and 'intersection'. An important addition to this assignment was that a starting spot of a truck would also be an ending spot and the other way around. Therefore, any column with source and sink type has been changed to just 'sourcesink' since it is the same.
Adding links in between two bridges	Sometimes there are two bridges nexst to each other in the dataframe. A link with length 0 has been added in between these bridges.

Replacing some specific intersection points	Some intersection points were located on the wrong location and, therefore, have been replaced manually.
Create overlap between intersections	In order to let intersections connect with each other, their ID must be identical. The system used for this phenomenon is based on the first intersection point of the two roads. A crossroad exists out of two roads with each an intersection point in the dataframe. The id of the first intersection point in the dataframe of a crossroad is the base for the id of the second intersection point of the crossroad. In order to
Combining links into one large link	Sometimes multiple links are located next to each other in the dataframe. These links are combined into one larger link to maintain the structure of link - bridge - link - bridge etc. (obviously, sometimes there is an instersection point in between.)

Table 2: Process of adding AADT to the input data

Subject	**Method of data cleaning**
Extract intersections and sourcesinks from the input file	For every road, the intersections and sourcesinks were extracted manually with the according road name, lat/lon and LRP name
Checking if intersections are there in real life	The lan/lon coordinates were copy pasted into Google Maps to see if the LRPs found to be intersections were really intersections in real life. The coordinates were adjusted as well if they were off
Find the corresponding AADT for the LRP in AADT-N-roads-overview.csv	The LRP names found are used to find the corresponding row in the AADT-N-roads-overview.csv file. Not all LRPs are in this file. If the LRP name did not correspond, the last LRP in the range is used for the AADT value. This is the closest LRP to the real point and represents the AADT of the intersection the best as it is directly connected to this intersection
Using the file for creating vehicles	The file is used create-input-data-roads.py and added to the input data of the model. The AADT is used by SourceSinks to create vehicles according to the real world number of vehicles passing that point of the road

Table 3: Overview of transforming data vehicles LRP locations

Subject	**Method of data cleaning**
Extracting information from traffic htm files	For the N1, N2, and N side roads, in the RMMS data package, there are separate files 'N.traffic.htm' accordingly which contain data on the number of trucks passing between two LRP points. Since the assignment only is about criticality and vulnerability regarding trucks, Heavy, Medium, and Small Trucks should only be accounted for. The traffic is measured as AADT, which is the total volume of vehicle traffic in both directions on a section of road for a year divided by 365 days. First, all the different traffic.htm files are copied into one CSV file manually to have one file which contains the necessary data on N1, N2, and N side roads. This has been done since this was expected to be faster than importing the separate htm files into python.
Cleaning the CSV file	The csv file has been cleaned manually by removing the columns which were not needed and rows that contain no values or inappropriate values.

Connecting the available information on trucks to the main data file	A difficulty with the information in the newly created CSV file is that it contains vehicle data between two random LRP points. This means that for instance, it is not clear exactly how much trucks drive between one sourcesink to an intersection when just between two points in between this road section there is data available on how many trucks have passed in a certain time frame. It might look like that if between two sourcesinks there is a road-section which contains the number of vehicles that it must be clear how many vehicles have passed that section. However, unfortunately, there are many instances that this cannot be assumed and is also not the case after checking the data.
First Consideration	Firstly, when a truck gets to an intersection it might go to another destination, and is not clear what path the trucks exactly are taking. The solution for our case to solve this is by replacing all intersections as sourcesinks. The assignment is to find out the critical bridges and thus it is not important which exact route the trucks are taking but how many trucks are passing a bridge in a time frame. Therefore, every intersection can be modeled as a sourcesink to represent much more accurately how many trucks are passing road section. This also means that NetworkX,which was added to generate random truck routes and get intersections to work, has been removed. The model for assignment 4 does not need this anymore since intersections are removed and random truck data is replaced by actual data. In the model description, this is further examined.

Second Consideration	Second, z roads and small side roads were removed from the data however, trucks could very well have used these roads and thus it is unclear how many trucks have crossed certain points in the network. A partial solution to getting the data as accurate as possible is to use the truck data between LRP's which are very close to sourcesinks. In this case, the chances are the lowest that they have been diverted to other roads.
Third Consideration	Third, it is not clear if these trucks were incoming or outgoing vehicles. Driving times in reality might be different since bridges might have a left and a right side. However, as the data has already been cleaned without making a distinction between left or right for bridges it is assumed that the direction of the truck is evenly divided. Thus if 100 trucks pass through 2 LRP points, 50 trucks go one direction and another 50 trucks the other way around.
Creation of data file for trucks LRP passing	In order to find out how many trucks have passed through every different intersection and sourcesink, manually all different road sections are implemented in a new file. For instance, from one sourcesink to an intersection or even from one intersection to another intersection. Then the CSV with the truck data is used and checked on which road the LRP's are and if the LRP's are between the two LRP's from the intersections/sourcesinks. A new CSV has been made which has data on the following columns: origin, lat, long, destination, lat, long, and number$_t rucks$.

4 Model Structure

4.1 File structure

There are seven files in the model structure responsible for different tasks. How these processes are linked and what every file produces is shown in figure 2 on the next page. The difference between model-run.py and model-viz.py requires elaboration since the process is similar but has important differences. The files both run the model with a different setup. Model-viz visualizes the Mesa model with a single manually set scenario of bridge conditions, which will be further explained in the next section 4.2. Model-viz requires a manual start and stop in the external web browser visualization. Model-run is able to run multiple scenarios and of each scenario multiple replications once the file is run and stops when the number of scenarios and replications are reached. The input for the files is the same produced by the files model.py, components.py, and create-input-data-roads.py. The process for create-input-data-roads.py was explained in section 3. The components.py and model.py file will be elaborated in the next section 4.2. The Simulator.py file allows for multi-threading or single-threading to speed up the run time of the model and experiments. Running the model works exactly the same as explained in the model-run swimlane, whichh is why this whole process is not copied in the Simulator.py for clarity and overview of the diagram.

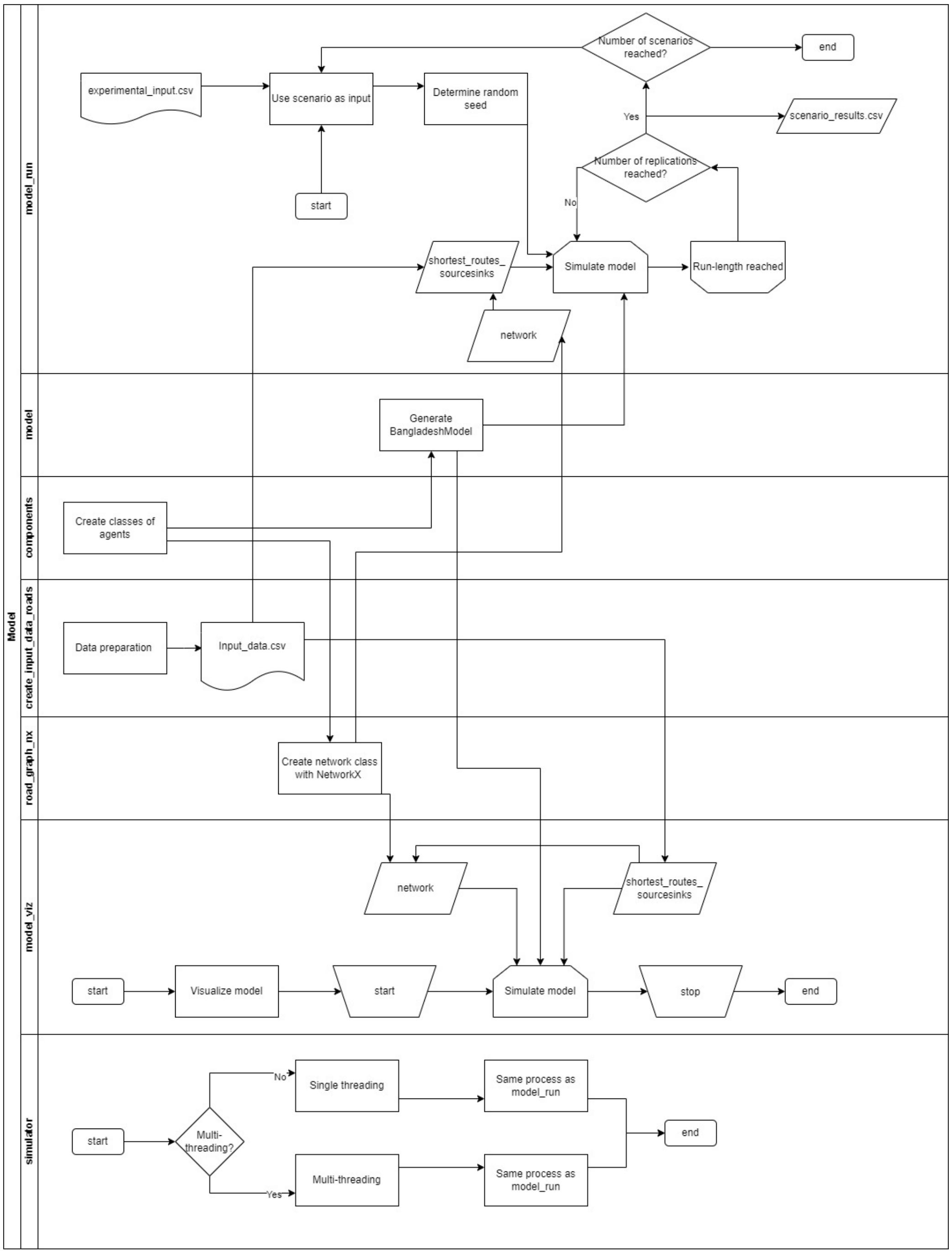


Figure 2: Swimlane diagram of Model Structure

Road-graph-nx.py is not used in this analysis since the paths created for the vehicles does not require the module Networkx to find the shortest path. However, the option to use the shortest paths as paths for the vehicles was added to the model for other analysis purposes.

4.2 Class structure

To create the different agents in the model produced by model.py, components.py builds different classes of infrastructure agents and vehicle agents. Figure 3 gives a clear overview of inheritance and how the classes are related to each other. The Model and Agents classes are used from the Mesa module. These classes show the necessary input and functions to build a Mesa model which does not add additional information to the figure. Their specific characteristics, functions, and values are left out of the diagram since they are "inherited" as a normal class would but they are specified in a more specific way in the BangladeshModel class and Infra and Vehicle classes. The classes from the Mesa module show the necessary input and functions to build a Mesa model which does not add additional information to the figure.

Moreover, the SourceSink class inherits methods and properties from both the Source and Sink class. This means that these agents are able to create and remove vehicles. This is one of the important elements that allow the model to create and remove vehicles that can travel in both ways.

The most complex component in the model is the Bridge class. The Bridge class is what creates delay time in the model based on the condition. Links or roads do not create delays. This condition (A, B, C or D) is gathered from the data frame per bridge. The idea is that condition A is the best state of a bridge and condition D is the worst. The bridges have a chance of breaking down according to this condition. The vehicles are affected by this delay time and have to wait accordingly at a bridge before they can travel across.

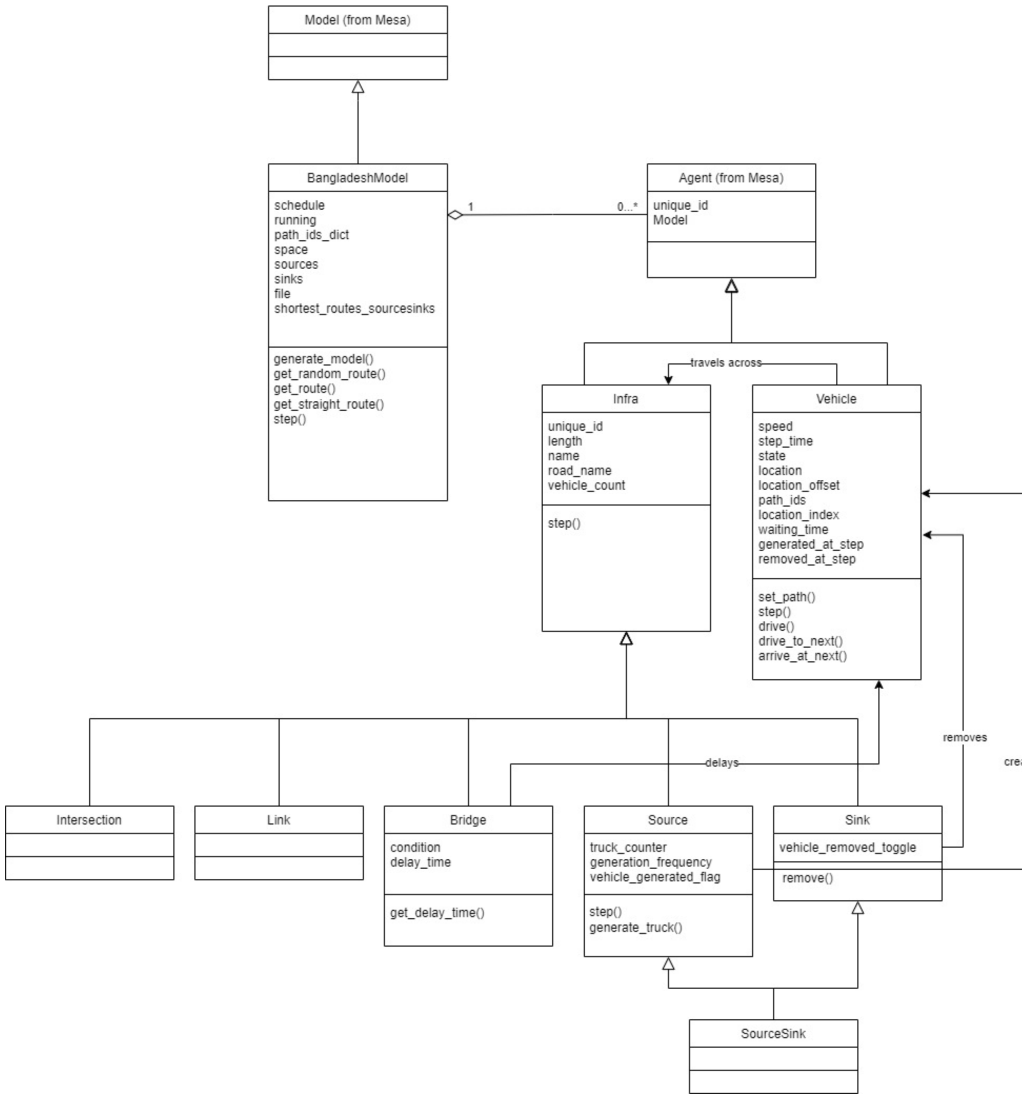


Figure 3: Class Diagram

5 Experimental setup

In the model are scenario implemented in order to analyse the behaviour of the model under different circumstances. The different scenarios are focussed on the chance of breaking down of a bridge. Each bridge has got a condition and each condition will have a certain chance of breaking down. In order to analyze the network and show the impact of bridges breaking down, different scenarios are introduced with replications to test the model.

5.1 Setup scenarios

The percentages that give the chance of a bridge breaking down per category of condition are displayed in Table 4 going from best to worst. This creates a different version of the network for every replication of a scenario, allowing for multiple replications to run and comparisons between them.

Table 4: Chance of break down per type of bridge per scenario

Scenario	Cat A	Cat B	Cat C	Cat D
1	0%	0%	0%	5%
2	0%	0%	0%	10%
3	0%	0%	5%	10%
4	0%	0%	10%	20%
5	0%	5%	10%	20%
6	0%	10%	20%	40%
7	5%	10%	20%	40%
8	10%	20%	40%	80%

Every bridge is either 'broken' or not after the scenarios are run. The delay time is assigned to every bridge according to the length of the bridge. For every range in length, a different stochastic distribution is used to determine the number of minutes every bridge delays the vehicles to cross. The distributions are shown in Table 5.

Table 5: Delay time per type of bridge

Bridge length	Delay time for a truck
under 10 meter	Uniform(10, 20) minutes
Between 10 and 50 m	Uniform(15, 60) minutes
Between 50 and 200 m	Uniform(45, 90) minutes
Over 200 m	Triangular(1, 2, 4) hours

The delay time will eventually function as waiting time for the vehicles. Bridges breaking down according to different scenarios and running the model based on these scenarios is described in the next section.

5.2 Iterative analysis

Each scenario described above will make use of 10 replications in order to minimize the effect of probability distributions on the results and, therefore, in order to generate more reliable results. Obviously, it is important to create a proper working model, but after completing the model it is important to analyse the behaviour of the model precisely. There are two ways to make sure that the behaviour of the model is analysed properly: by creating enough and sharp scenarios and by using enough replications of each scenario in order to improve the reliability of the output of a scenario.

In this study it was not possible to make lots of replications for all the scenarios, since the running time of the model would be too large. Therefore, only scenario 8 is picked for the iterative analysis. The iterative analysis implies a set of 30 iterations instead of 10 iterations for the model run. As stated before, the improvement of the number of iterations will decrease the effect of chance probabilities within the model. Scenario 8 is selected for the iterative analysis, since this scenario contains the largest chance for all types of bridges to break down. Therefore, this scenario will create the most number of bridges to collapse and, therefore, is most likely to create the highest delay time. That is why the difference in output by applying 30 replications instead of 10 replications is the most visual for scenario 8 and, therefore, has been applied for scenario 8.

6 Results

The results chapter is divided into three subsections: the first subsection is information on the most critical bridges under different scenarios. The second subsection shows the results for the vulnerability analysis. The third section focuses on the iterative analysis done with scenario 8.

6.1 General results

In order to have a first look on the different behaviours of the model within the implemented scenarios, box plots are created. Box plots are a method in which data is visualized in such a way that the spread of data points is very clear. The box plots created are made on the delay time per bridge in the model. However, it must be noted that bridges which are not broken have no impact on the box plots.

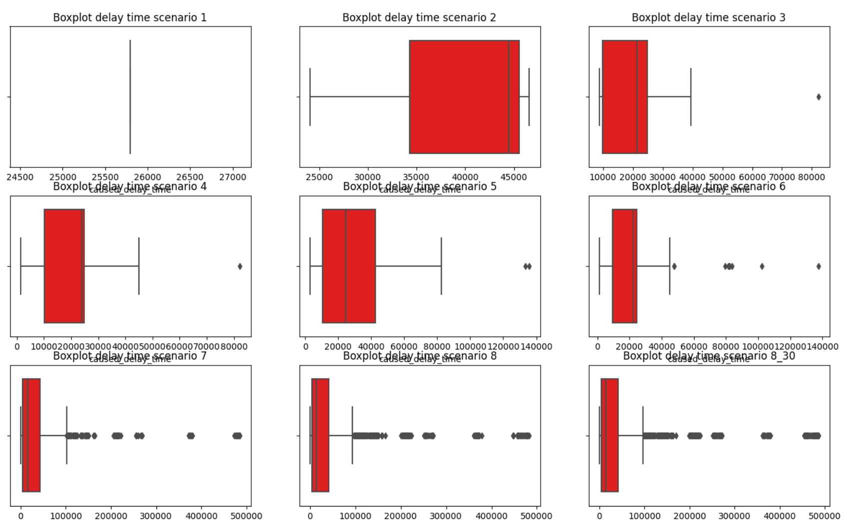


Figure 4: Box plots of all the scenarios implemented

The first thing that stands out is the box plot for the first scenario, since it is just a thin line. this can be explained by the fact that this scenario contains only 1 bridge that collapses, due to the fact that there is a very low possibility for just the bridges with condition D to collapse. Furthermore, the box plots show the behaviour as expected from the structure of the scenarios: the median of the box plots increases over the scenarios, just as the number of outliers. These outliers are formed by bridges with very high delay time when they break down due to their length. The chance of breaking down is relatively low, since lots of these long bridges are of condition A or B, but the last couple of scenarios contain bridges with condition A and B to break down as well. Due to these outliers, the spread of data gets larger across the scenarios.

6.2 Criticality analysis individual bridge results

The goal of the study is to identify the most vulnerable and critical bridges in Bangladesh. Different scenarios are important to test with the model since Bangladesh's climate disasters may have relatively small or enormous consequences for its infrastructure. In order to assess the results, box plots have been made for visualization. Box plots are useful to see the spread of data points (Insightoriel, 2022). Finally, before interpreting the scenarios it must be stated that the model uses stochastic functions in order to calculate the number of broken bridges. Therefore, the scenarios differ for each run of the model. In order to be able to investigate the distribution of the output, more replications are applied for each scenario. In this research, there are 10 replications applied. In the following tables, the top 5 most critical bridges are identified for all 8 scenarios. For these bridges, the number of vehicles, location, and total delay time are given.

Table 6: Scenario 1 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	982	680	25797	"N1, LRP293a - LRPE"	92.1413608,21.242333

Table 7: Scenario 2 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	904	671	45532	”N1, LRP293a - LRPE”	92.0773327,21.3731938
2	982	641	24052	”N1, LRP293a - LRPE”	92.1413608,21.242333

Table 8: Scenario 3 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	2189	5470	82297	”N2, LRP117b - LRP146b”	91.3474441,24.1498886
2	904	595	39429	”N1, LRP293a - LRPE”	92.0773327,21.3731938
3	852	675	25974	”N1, LRP293a - LRPE”	92.0531944,21.4094167
4	868	672	24769	”N1, LRP293a - LRPE”	92.068166,21.393333
5	982	652	24464	”N1, LRP293a - LRPE”	92.1413608,21.242333

Table 9: Scenario 4 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	2189	5470	82243	”N2, LRP117b - LRP146b”	91.3474441,24.1498886
2	904	647	43519	”N1, LRP293a - LRPE”	92.0773327,21.3731938
3	852	668	25683	”N1, LRP293a - LRPE”	92.0531944,21.4094167
4	1020	656	24685	”N1, LRP293a - LRPE”	92.1545833,21.1698049
5	2420	1641	24646	”N2, LRP146b - LRP191b”	91.5720456,24.45314813

Table 10: Scenario 5 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	482	2004	134678	"N1, LRP184a - LRP260c"	91.897333,22.3015833
2	2272	5495	82577	"N2, LRP117b - LRP146b"	91.3808604,24.2633333
3	2189	5459	81700	"N2, LRP117b - LRP146b"	91.3474441,24.1498886
4	513	3176	47994	"N1, LRP260c - LRP293a"	92.0075556,22.2600833
5	938	666	44909	"N1, LRP293a - LRPE"	92.0787771,21.3297771

Table 11: Scenario 6 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	482	2031	137649	"N1, LRP184a - LRP260c"	91.897333,22.3015833
2	176	6828	102465	"N1, LRP084a - LRP184a"	91.3098056,23.248333
3	2189	5485	82415	"N2, LRP117b - LRP146b"	91.3474441,24.1498886
4	2271	5446	81498	"N2, LRP117b - LRP146b"	91.3801382,24.2620833
5	2792	2125	79518	"N2, LRP239a - LRPE"	92.1056389,25.1546944

Table 12: Scenario 7 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	37	7189	485062	"N1, LRP009a - LRP084a"	90.5970274,23.6341108
2	17	7197	484338	"N1, LRP009a - LRP084a"	90.5410552,23.6922774
3	39	7171	483648	"N1, LRP009a - LRP084a"	90.6042222,23.6262771
4	55	7124	481351	"N1, LRP009a - LRP084a"	90.6381386,23.5825827
5	81	7065	477216	"N1, LRP009a - LRP084a"	90.7173604,23.5315271

Table 13: Scenario 8 Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	37	7155	483473	"N1, LRP009a - LRP084a"	90.5970274,23.6341108
2	101	7124	480250	"N1, LRP009a - LRP084a"	90.91525,23.50575
3	113	7078	479179	"N1, LRP009a - LRP084a	91.0404441,23.481166
4	63	7089	478137	"N1, LRP009a - LRP084a"	90.6542222,23.5604167
5	81	7051	476590	"N1, LRP009a - LRP084a"	90.7173604,23.5315271

The tables showcase the bridges which caused the highest delay time. First, it is noticeable that several bridges reoccur in different scenarios multiple times: 982, 904, 2189, 852, 482, 37, and 81 (by ID number). Especially bridge 2189 is interesting in scenarios 3, 4, 5, and 6. In these scenarios Category C bridges have a chance of breaking down. Moreover, the impact of category C and category B bridges breaking down has a relatively small difference between them in these scenarios. This is an exception that will be elaborated on further in this section.

The fact that there are some reoccurring bridges gives the sense that the data is not that random and that advice to invest in certain vulnerable infrastructure parts is robust. For bridge 2189, the total delay time is the following: 82297, 82243, 81700, and 82415. These total delay times are only slightly different and this further increases the sense that the stochastic breaking down of bridges per condition in the model and the input does not result in random output. Due to the low number of replications, this is not a valid claim. Therefore, 30 replications have been done for scenario 8 to test whether the results are actually relatively robust.

Another insight is that the category of bridges and the assigned vulnerability to this category has many effects on the results. When a new Category is added to a scenario, different bridges turn out to be critical. For instance, scenarios 6 and 7 are the same with only the addition of Category A bridges being vulnerable. The results show a completely different list of the top 5 most critical bridges. This also gives a further implication that category A bridges score higher on criticality than other bridges. Moreover, this holds for the other categories as well where Category C bridges have more impact than Category D and so on. This has to do with the fact that delay times increase faster for longer bridges than for small bridges. It also has to do with the number of vehicles crossing a bridge since the top 5 of scenario 8 bridges show much more trucks crossing than the top 5 of scenario 1 bridges.

In the scenarios where the highest categories of bridges do have a chance of breaking down, the delay time is much higher than in the other scenarios. This implies that higher-category bridges are more critical to the road network since they cause more delay once broken. Moreover, higher-category bridges are often longer which will result in more delay time once the bridge is broken.

The 5 most critical bridges in order are the following based on this analysis: 37, 81, 482, 2189, and 904. This is based on the reoccurring of bridges in different scenarios, the fact that they mostly are the top 1 vulnerable in their respective scenario, and the total delay times they cause for trucks.

All in all, it can be concluded that the vulnerability of bridges is much dependent on different scenarios and thus chances of breaking down. The criticality of bridges is largely dependent on the type of bridge rather than the number of vehicles. However, more research must be done to verify this statement.

6.3 Criticality analysis road segments results

As mentioned before, this part of this study focuses on the criticality of the model. This subsection contains the results of the criticality analysis results for larger road segments instead of individual bridges. By looking at a road segment instead of individual bridges, one can establish whether bridges within the same region have to be reconstructed. The criticality analysis for road segments has been done for both scenario 4 and scenario 8. This choice is made, since these scenarios offer insights into different kinds of circumstances: a scenario in which the storm is strong, but will not cause major destruction and a scenario in which a storm would cause major destruction across the country. Based on the expected future, different decisions are related to the order of reconstruction. It must be noted that the number of broken bridges is based on the top 100 bridges that cause delay times in the model. In this way, the 100 most important individual bridges are part of this analysis. The results of the analysis are shown in tables 14 and 15.

Table 14: Top 4 road segment with most bridges broken (scenario 4)

road segment	Number of broken bridges
N1, LRP293a - LRPE	19
N2, LRP117b - LRP146b	1
N2, LRP146b - LRP191b	1
N2, LRP191b - LRP228c	1

Table 15: Top 5 road segments with most bridges broken (scenario 8)

road segment	Number of broken bridges
N1, LRP009a - LRP084a	21
N1, LRP084a - LRP184a	15
N2, LRP012a - LRP086a	14
N102, LRPS - LRPE	11
N2, LRP117b - LRP146b	9

One thing that is remarkable is the fact that there are only 4 road segments in scenario 4 that contain broken bridges. This is due to the fact that there are only 23 bridges that will break down in that scenario. These bridges are located in just 4 road segments and, therefore, only 4 road segments are part of this analysis. Table **??** shows a top 5 road segments since there are enough road segments with broken bridges. Furthermore, it can be seen that the road segment N1, LRP293a - LRPE contains many bridges being broken down in scenario 4, but the broken bridges in this road segment do not cause very high delay times in scenarios 8, since this road segment is not in the top 5. It must be noted that this does not mean that the number of broken bridges on this road segment is less in scenario 8 than in scenario 4. It only means that the number of bridges that cause top 100 delay time in scenario 8 is less than the number of bridges that cause relatively high delay times in scenario 4. Furthermore, as can be seen in table 14, for a scenario in which Bangladesh has to do with not very heavy storms it is very important to focus on the bridges on the N1, LRP293a-LRPE. Finally, it is remarkable that a large amount of the top 100 bridges are located in only the top 5 road segment parts. 70 of the 100 bridges which cause massive delay times are located on these 5 road segments.

6.4 Vulnerability analysis results

Besides criticality, it is important to analyze the vulnerability of the model as well. The vulnerability of the network represents whether the network is likely to break down or not. During the vulnerability analysis, the condition of the 5 bridges which cause the most delay time is shown. Via this method, it can be seen whether the bridges which cause major delay time are of good quality or not. Just like the criticality analysis for road segments, this analysis uses scenario 4 and scenario 8, representing a mediocre and a very heavy storm. The results of the analysis are shown in table 16 and 17.

Table 16: Conditions of top 5 broken bridges scenario 4

Bridge ID	condition	Road segment	lat,lon
2189	C	N2, LRP117b - LRP146b	24.1498886, 91.3474441
904	D	N1, LRP293a - LRPE	21.3731938, 92.0773327
852	C	N1, LRP293a - LRPE	21.4094167, 92.0531944
1020	C	N1, LRP293a - LRPE	21.1698049, 92.1545833
2420	C	N2, LRP146b - LRP191b	24.453148133, 91.5720456

Table 17: Conditions of top 5 broken bridges scenario 8

Bridge ID	condition	Road segment	lat,lon
17	A	LRP009a - LRP084a	23.6922774, 90.5410552
37	A	N1, LRP009a - LRP084a	23.6341108, 90.5970274
55	A	LRP009a - LRP084a	23.5825827, 90.6381386
63	A	LRP009a - LRP084a	23.5604167, 90.6542222
97	A	LRP084a - LRP184a	23.5133056, 90.8703049

The most important thing noticeable is the difference in the type of condition between scenario 4 and scenario 8. In scenario 4, only bridges with conditions C and D are in danger of breaking down. The delay time caused by the broken bridges is not so much due to the fact that the broken bridge is a large bridge, but mostly due to the fact that the broken bridge is a bridge with a high traffic flow. Therefore, the small amount of delay time generated for each truck will still cause relatively high delay times. In scenario 8 all the bridges in the top 5 are of condition A. This is be as expected since this scenario contains bridges of class A to break down. Bridges of class A are mostly located in very dense locations and are often longer bridges. Therefore, longer delay times are generated per truck and the total number of passing trucks is high.

6.5 Iterative analysis results

As stated before, the iterative analysis checks whether the output of a scenario differs when the number of iterations is adjusted. In this study, scenario 8 is selected for the iterative analysis, since this scenario contains the most information of all scenarios. The result of the analysis is shown in tables 18 and 19.

Table 18: Scenario 8, 10 replications: Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	37	7155	483473	"N1, LRP009a - LRP084a"	90.5970274,23.6341108
2	101	7124	480250	"N1, LRP009a - LRP084a"	90.91525,23.50575
3	113	7078	479179	"N1, LRP009a - LRP084a	91.0404441,23.481166
4	63	7089	478137	"N1, LRP009a - LRP084a"	90.6542222,23.5604167
5	81	7051	476590	"N1, LRP009a - LRP084a"	90.7173604,23.5315271

Table 19: Scenario 8, 30 replications: Top 5 bridges

Top	ID	Vehicles	Delay time	Road part	Lat, Lon
1	17	7191	485167,77	"N1, LRP009a - LRP084a"	90.5410552,23.6922774
2	37	7169	484459,87	"N1, LRP009a - LRP084a"	90.5970274,23.6341108
3	55	7142	481880,20	"N1, LRP009a - LRP084a	90.6381386,23.5825827
4	63	7120	480665,67	"N1, LRP009a - LRP084a"	90.6542222,23.5604167
5	97	7105	480496,59	"N1, LRP009a - LRP084a"	90.8703049,23.5133056

In stochastic simulation experiments, at least thirty replications per scenario are the minimum, indicated by the central limit theorem, stating that this number is the least required to ensure normally distributed results (Frost, 2022). Due to time limitations, this assignment only managed an insufficient ten replications. However, to still get an idea of how accurate runs are and how much the results are affected by randomness in a stochastic simulation, for scenario 8 there a run has been done for 30 replications. In 19, the top 5 most critical bridges can be seen. Additionally, in 18, the simulation with 10 replications can be seen again. Two bridges occur in both tables which indicates that the critical bridges are relatively random based on the simulation. However, it can be seen that these bridges all are on the same road part. This indicates that instead of focusing on just a few bridges, parts of the network should be targeted where vulnerable bridges are located. Furthermore, all the bridges are category A bridges which indicates that the type of category of bridges is not affected by different simulations.

7 Discussion

This model aims to make an accurate representation of the road network of Bangladesh to analyze the consequences of bridges breaking down for logistic traffic. To do that, the hourly traffic passing an LRP (AADT) is used to create vehicles and to simulate a real-life traffic flow on the most important roads of Bangladesh. An assumption had to be made for selecting the AADT for LRPs not present in the AADT-N-roads-overview file. The last LRP in the sequence between two LRPs was selected to represent the AADT. However, there are several other ways this assumption could have been made. One could argue that the road segment has as much traffic passing as the lowest or highest AADT. Moreover, another LRP, for instance the first in the segment, could have been selected to represent the AADT. This would structurally change the traffic flow simulated and would influence results.

A potential limitation of the model is in running time for the experimental setup. Namely, as a follow-up to the above, it would be logical to run more replications and scenarios. The current experimental setup (with only 8 scenarios and 10 replications per scenario) has a running time of about 1 hour. In order to improve the results, as mentioned above, the number of scenarios and replications must be increased which will increase the total running time. It would be possible to execute the above additions and still have a reasonable running time. However, the model only contains the N1, N2, and a few side roads. When the complete road network of Bangladesh is implemented, the running time will probably be too long. It is important to reflect on this. This means that other experimental design methods should be used to obtain accurate results. A concept that can be used is space-filling designs (Sanchez & Wan, 2021). Instead of running every possible scenario (full-factorial design) or only a few constructed manually, a systematic selection can be used. Currently, in the model, the scenarios assign bridges different 'caused-delay' parameters, based on five different scenarios. For accurate scenario analysis, many more scenarios with different parameters must be tested. Sanchez and Wan (2021) describe this as a pitfall when a handful of scenarios is chosen which may not address fundamental questions. Space-filling designs could again be useful. It is important to consider this in further research.

While running the experimental setup, every scenario is run for 10 replications. To properly remove stochastic influences and make valid statements on which bridges are critical and vulnerable, more replications are necessary. Due to time limitations, this is not incorporated in this report.

Another limitation is found regarding the merging of the files. In the data generation process for the model, the road and BMMS (bridge) files were merged. The method used for this is 'right-merge', which resulted in the road file being the 'main file' and based on overlapping road and LRP, the records from BMMS were merged with those of the road file. The road file fitted the required data format better and was more complete. Moreover, the roads file had fewer double identifiers (LRPs) for different objects in comparison with the BMMS file. This way of merging makes sense since the structure is built up by links (road parts) and bridges in the roads file. More information was needed from the BMMS file on the bridges breaking down which was added to the roads file. Intuitively, this is also the correct way, because bridges are on (or nearby) roads. If we merged the other way around (left-merge), some parts of the road would not be taken into account. The road would simply stop at some point or not be connected, which is not an option. An inner merge would not work in this case as the road parts in the road file would not be taken into account. This is because only the bridges in the road file are merged with those in the BMMS if the LRPs are consistently used in both files. This leaves only bridges in the data frame. An outer merge is not chosen, occasionally the same bridge name does not have the same LRP in both files. With an outer merge, both these bridges (the same bridges, but with different LRP) will be included in the data frame used for the model, which is incorrect.

Therefore, we used a right-merge. However, we did not specify that only the left or right bridges should be taken into account from the BMMS file if the LRPs are exactly the same and overlap with a record in the roads file. Instead, the first record in the BMMS file with the overlapping LRP for the bridge is merged into the data frame and used in the model. In general, this fits the assumption of vehicles driving two ways better because sometimes the right bridge is pulled from the data frame and sometimes the left bridge is. If trucks drive in both directions, this is the way to go. This is the case since starting and ending nodes are one and the same in this model. However, separate nodes could be made for left and right bridges since in reality when for instance, the left bridge breaks down, the right lane traffic can continue driving. This would be an improvement that does require separate nodes for bridges left and right and a distinction for trucks that the direction implies whether it can use a certain node. This change would greatly complicate the model.

To accurately conclude the impact of broken bridges on the road network in Bangladesh, more research is necessary. However, this report provides an insightful starting point.

8 Conclusion

Bangladesh's infrastructure has been found to be vulnerable to natural disasters where the scale of a disaster has a great impact on the infrastructure and the accessibility of truck transport. The different scenarios have reflected the impact of a natural disaster. For instance, scenario 4 can be seen as mediocre and scenario 8 is a major natural disaster. The results have shown that the type of disaster has large impact on the delay times of trucks due to the different chances, quantity, and quality of bridges breaking down. Therefore, several bridges have been identified which are the most critical and should be invested in to help Bangladesh cope with infrastructure challenges. The 5 most critical bridges on id in order are the following based on this analysis: 37, 81, 482, 2189, and 904. This is based on the reoccurring of bridges in different scenarios, the fact that they mostly are the top 1 vulnerable in their respective scenario, and the total delay times they cause for trucks. This limited number is bound by the budget available and any further investment in the critical parts of the infrastructure can be focused on the next most critical and vulnerable parts of the infrastructure.

The research is limited by the fact that only a small part of the network of Bangladesh is analyzed, and the stochastic behavior with the limited simulation runs done due to the limited time available and the computing power capabilities. Furthermore, in the data-cleaning process, several assumptions have been made and some could be improved by further analysis and more time. Besides, the critical and vulnerable bridges could a more extensive analysis of sections of the network give more insights into which parts of the network are actually critical. This may lead to another assessment of the critical infrastructure of Bangladesh. Further research can extend this model with adding more of the Bangladesh roads to the model and simulate the road network more extensively.

References

Berdica, K. (2002, Apr). An introduction to road vulnerability: What has been done, is done and should be done. *Transport Policy*, *9*(2), 117–127. doi: 10.1016/s0967-070x(02)00011-2

Erath, A., Birdsall, J., Axhausen, K. W., & Hajdin, R. (2009, Jan). Vulnerability assessment methodology for swiss road network. *Transportation Research Record: Journal of the Transportation Research Board*, *2137*(1), 118–126. doi: 10.3141/2137-13

Frost, J. (2022, 2). *Central Limit Theorem Explained.* . (Online; accessed 10 March 2023)

García-Palomares, J. C., Gutiérrez, J., Martín, J. C., & Moya-Gómez, B. (2018, Apr). An analysis of the spanish high capacity road network criticality. *Transportation*, *45*(4), 1139–1159. doi: 10.1007/s11116-018-9877-4

Insightoriel. (2022). *What Is Boxplot — Box And Whisker Plot — 5 Advantages Of Boxplot — Create Boxplot In Excel.* . (Online; accessed 10 March 2023)

Jafino, B. A., Kwakkel, J., & Verbraeck, A. (2019, Jul). Transport network criticality metrics: A comparative analysis and a guideline for selection. *Transport Reviews*, *40*(2), 241–264. doi: 10.1080/01441647.2019.1703843

Jenelius, E., & Mattsson, L.-G. (2015, Jan). Road network vulnerability analysis: Conceptualization, implementation and application. *Computers, Environment and Urban Systems*, *49*, 136–147. doi: 10.1016/j.compenvurbsys.2014.02.003

Kaplan, S., & Garrick, B. J. (1981, Mar). On the quantitative definition of risk. *Risk Analysis*, *1*(1), 11–27. doi: 10.1111/j.1539-6924.1981.tb01350.x

Klabunde, W. F., A. (2016). Decision-making in agent-based models of migration: State of the art and challenges. *Eur J Population*. doi: https://doi.org/10.1007/s10680-015-9362-0

Luathep, P., Sumalee, A., Ho, H. W., & Kurauchi, F. (2011, Jun). Large-scale road network vulnerability analysis: A sensitivity analysis based approach. *Transportation*, *38*(5), 799–817. doi: 10.1007/s11116-011-9350-0

Macal, C., & North, M. (2005). Tutorial on agent-based modeling and simulation. In *Proceedings of the winter simulation conference, 2005.* (p. 14 pp.-). doi: 10.1109/WSC.2005.1574234

Sanchez, S. M., & Wan, H. (2021). Work smarter, not harder: A tutorial on designing and conducting simulation experiments.. doi: 10.1109/WSC.2015.7408296