EPA1352. Advanced Simulation

Assignment 2. Building Components for Data-Driven Simulation

Group 17

Ivan Temme (4955196)

Hidde Scheuer (4607325)

Philip Mueller (5809703)

Madalin Simion (5838363)

Nachiket Kondhalkar (5833884)

1. Introduction

This report aims to analyse the impact of bridge unavailability and maintenance on traffic travelling from Chittagong to Dhaka by creating a Mesa-model based on components. To reduce the influence of modeller bias (Keller & Hu 2019), a data-driven modelling approach was employed in this study.

2. Data preparation

The data preparation and cleaning consisted of multiple tasks that will be outlined in the following. In this report we will address and justify the major design choices that we have made. Data preparation resulted in a clean data set that combines data on roads and bridges and is ready to be fed into the simulation model. For the purpose of this assignment, all simulations have been done on the highway "N1".

2.1. Dataframe size and merging

The N1 Highway spans from Dhaka to Cox's Bazar (~463km). In this assignment, we are only interested in the section between Dhaka and Chittagong. Therefore, only the data for the first 241km (N1 starting at Dhaka) will be kept in the dataset. After km 241.1 the Chittagong area ends (according to the last linkage, entry 567 in *_roads3* file).

There are no missing points in the *_roads3* dataset and neither are there any duplicates. However, for the purposes of the assignment we need to know the quality of the bridge as well as its name. Hence, we also need to use the *BMMS_Overview* dataset to combine the points at which the road passes over a bridge. This is done using LRPs. However, the roads dataset has 1339 LRPs and the bridge dataset has 639. Of these, 415 bridges have LRPs that coincide with the LRPs on the road dataset. There are 224 bridges that have not been accounted for.

To avoid mistakes and simplify the Dataframe, the chainage was converted from km to the S.I. unit of meter. In addition, the length data from the *BMMS* file was combined with the chainage data from the roads file. After removing the part of the road from Chittagong to Cox's Bazar, the two data frames were merged based on the chainage value.

2.2. Dual bridges and duplicates

In the *BMMS* file some bridges are noted as 'L' and 'R', or 'left' and 'right', or are even duplicated. The distinction can be because there are different data inputs for the separate ways of travel, or because there are two different bridges (one older and one newer), each for a way of travel. Because in this model we are interested only in one-way travel, keeping both travel ways is unnecessary.

Therefore, in our code we have omitted the left side of the bridge as suggested in the assignment. Both sides of the bridge had complete data for the attributes we are interested in, such as condition or length and name, therefore the removal did not affect our model. This removal was performed by analysing the last few characters of the 'name' column for the bridge.

Even after removing the 'L' bridges, there were still some duplicates, with (almost) identical parameter values. Because the building year was different, we assumed that a new bridge was built next to the old one. As we expect the bridge of the latest construction year to be the most likely one to be still in place, we have removed all duplicates with an earlier construction year than the latest.

The dataset was then merged according to the LRPs and certain identifiers were added to the Dataframe according to the type of the road. A column *model_type* was added and 'bridge' was allotted to instances of points from the *BMMS_Overview* dataset and 'link' was allotted to every other point on the road. The data frame was then arranged in such a way that the start of the Dataframe would be in Chittagong and would be titled as 'source' and Dhaka would be the 'sink' for the vehicles. This would simulate the behaviour of the trucks as they picked up containers from the port at Chittagong and drove them up the N1 highway to Dhaka

3. Model Design

In order to simulate vehicles driving over the N1 road, modifications and extensions were made to various parts of the provided model. Changes to the *model.py* file are being discussed in chapter 3.1 and changes to the agent classes in the *components.py* file are being discussed in chapter 3.2. To run multiple iterations on various parameter setting of the model, some changes were made to the *model_run.py* and a new file, called *batch_run.py* was created, see chapter 3.3.

3.1. Changes to the Model

The model underwent several changes to provide the desired functionality. It now uses the *N1.csv* file as input to accurately model the correct road in the correct direction. In addition, four new inputs were added to represent the probability of bridges in each category breaking down. Bridge quality is categorized in four steps from A (good condition) to D (bad condition). These probabilities are then used by the Bridges class to determine which bridges are being modelled as broken down.

In this way we model destruction due to a natural disaster. Furthermore, four new functions were introduced as model reporters. The first function *compute_average_driving* calculates the average driving time for all trucks after a single model run. The second function *compute_worst_bridge* returns the name of the bridge with the highest total delay time. The third function c*ompute_worst_bridge_delay* returns the average delay time at the worst bridge. Finally, the fourth function *get_probs* provides the probabilities of bridges breaking down that were used in the model as an additional method of validation.

3.2. Changes to Agent Classes

Several (sub-)classes of agents underwent modifications in the model. In the Vehicle class, speed was changed to be set to 48km/h. The Sinks class now stores the time tick at which each truck is removed from the model, which will later be used to calculate driving times. The Bridges class underwent four changes. Firstly, their categories (A, B, C, or D) are now a parameter that is passed during initialization. Secondly, bridge instances can now break down based on a user-defined probability that is dependent on the bridge's condition. Thirdly, a new method, called *get_delay_time* was created. For each arriving truck the method computes a delay time, following a predefined probability distribution, based on the bridge's length. The delay time equals 0 for every bridge instance that is currently not modelled as collapsed. Lastly, Bridges now keep track of the number of trucks passed and the accumulated delay caused. This information will be used to calculate metrics later.

3.3. Batch Run

The *batchrunner.py* file was added to the project, which utilizes Mesa's batchrunner (https://mesa.readthedocs.io/en/stable/apis/batchrunner.html) to run all scenarios sequentially. Each scenario runs 10 iterations for a period of 5 days, which is equal to 52460 minutes. Here, one-minute equals one tick. A new vehicle spawns every five ticks. The list containing the probability of bridges breaking down for each category is passed as an argument to the batchrunner, along with the model reporters (see chapter 3.1 for the newly defined methos in model.py). The output is saved in one csv file for every scenario and in one combined file, called *scenarios.csv.*

4. Results

The *batchrunner.py* file contains scenarios 0 to 8 with varying probabilities of bridges breaking down for different condition-categories (A, B, C, D). After analysing the simulation outcomes of these scenarios, important findings have been summarized in several plots. These plots reveal that the model shows the anticipated behaviour. Driving time increases with the higher likelihood of bridge breakdowns.

In the following, scenario 0 will refer to the baseline scenario, in which no bridge is being modelled as collapsed. Table 2 shows the probability functions that are given to model delay for an individual vehicle approaching a bridge that is modelled as collapsed. As can be seen, the delay time increases with bridge length. The spread of randomly selected delay times is smaller for bridges with a length of more than 200 metres, since here a triangular distribution is being used, instead of a uniform distribution.

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

Table 1: probabilities of bridges breaking down, depending on their category and the modelled scenario. The table does not include the baseline scenario 0, in which all probabilities are set to 0.

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

Table 2: probability functions for modelling the delay at a collapsed bridge.

4.1. Driving and Delay Times

As can be seen in Figure 1, the average driving time is increasing with each scenario. The spread between minimum and maximum driving time, as well as the standard deviation are increasing too. The averages standard deviations of delay times are increasing too with each scenario. For more metrics, consult table 3 in the Appendix. Delay time influences driving time – the higher the total delay time of all vehicles in one model run, the higher the total driving time in the same model run.

The baseline scenario shows a driving time of 585 hours. That time only starts to increase in scenario 3. Hence, no bridge has collapsed in scenarios 1 and 2, in which only category d bridges are modelled for collapse. The provided dataset contains no category d bridge.

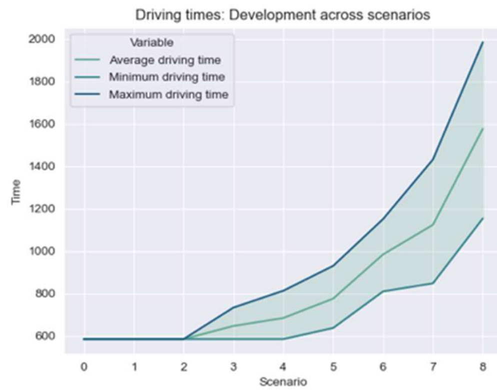

Figure 1: minimum, mean, and maximum driving times for each scenario.

The average driving time in scenario 8 is roughly between two and four times higher (roughly 3 times higher on average), compared to the baseline scenario. The only moderate increase in driving time can again be explained by the lack of category D bridges, and the relatively low occurrence of category C bridges (see 'Limitations'). Driving time is thus mostly influenced by category A and B bridges whose probabilities of collapsing are only spanning from 0% to 10% and 0% to 20% between scenarios 0 and 8.

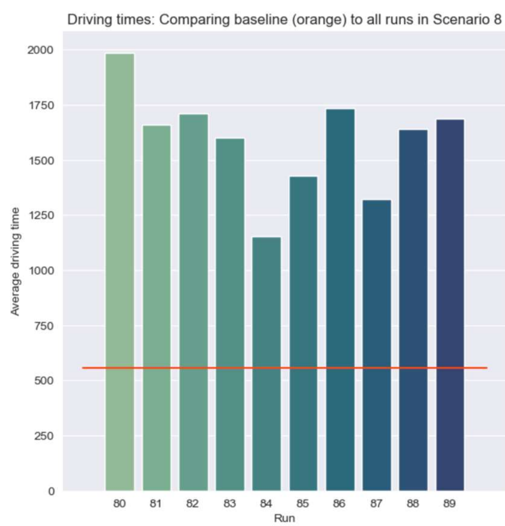

Figure 2: average driving time in each model run from scenario 8, compared to the driving time of all vehicles in the baseline scenario (orange line).

4.2. Bridges

Table 3 shows the 5 bridges that collapsed most frequently across all simulations. All 5 bridges are of category c. The bridges with rankings 1, 2, and 4 are showing a considerably higher delay time than the remaining two. The distribution of the delay times across these three bridges and all simulations can be seen in Figure 3. These bridges are longer than 200 metres. Their modelled delay time is thus based on the same probability distribution (see Table 2). How strongly ultimate delay time is being influenced by the shape of these probability distributions, can be seen in Figure 4 in the appendix. Muhuri Bridge falls into the length category between the ones with higher and lower delay times.

Ranking	Bridges name	Number of times damaged
1	KANCHPUR PC GIRDER BRIDGE	10
2	MEGHNA BRIDGE	8
3	Ginlatoly	5
4	Daud Kandi Bridge	5
5	BORO KUMIRA BRIDGE	4

Table 3: Most frequently collapsed bridges and number of times the bridge has collapsed.

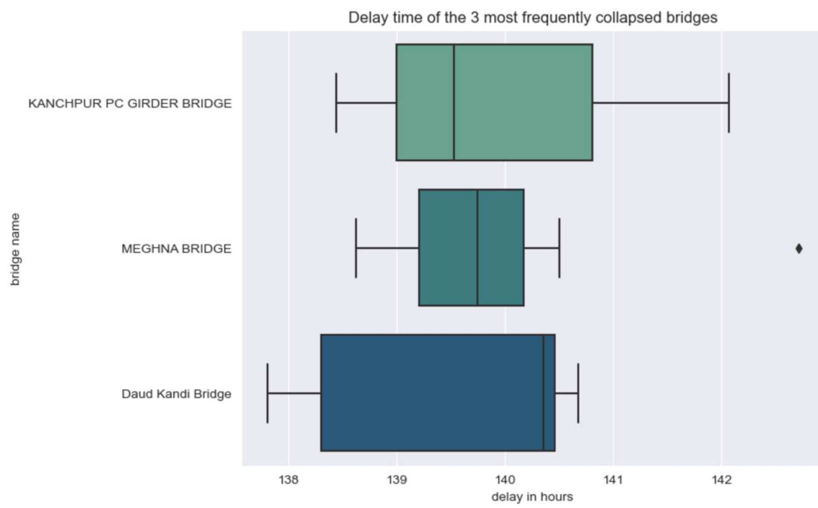

Figure 3: Boxplots of delay caused by the worst three bridges across all scenarios.

Based on the number of collapses and the resulting total delay time, Table 3 shows the five bridges that require most repairs investments by the government. It is worth noting, however, that the input parameter It is worth noting, however, that the number of collapses and delay times are heavily depending on the input parameters to the simulation. For instance, wow strongly the resulting delay time is being influenced by the shape of delay probability distributions, can be seen in Figure 4. Muhuri Bridge falls into the length category between the bridges with higher and the ones with lower delay times. Before drawing conclusions, the external validity of the employed probability delay functions should be checked. It is possible that delay times are depending on further factors, apart and independent from bridge length. It might be furthermore necessary to decrease bin sizes of the delay probability functions.

5. Limitations

There are multiple avenues for improving the model. Firstly, since the provided dataset lacks category D bridges, the impact of breakdowns in these bridges is absent from the findings, resulting in the worst bridge delay only appearing in scenario 3. Furthermore, out of 210 bridges, only 21 were category C, which implies that these bridges were not significant due to their underrepresentation.

Another limitation of the analysis is that it excluded 224 bridges from the dataset due to missing LRPs. Moreover, the approach of eliminating 'Left' side bridges may result in the exclusion of potential bridges that do not have a corresponding counterpart on the 'Right' side in the dataset in the future. This method needs to be monitored when scaling the model in the future.

The small sample size of only 10 iterations per scenario is the reason for the abrupt and jarring graph shown in *Figure 2*. If the number of iterations had been increased, the transitions between scenarios would have been smoother.

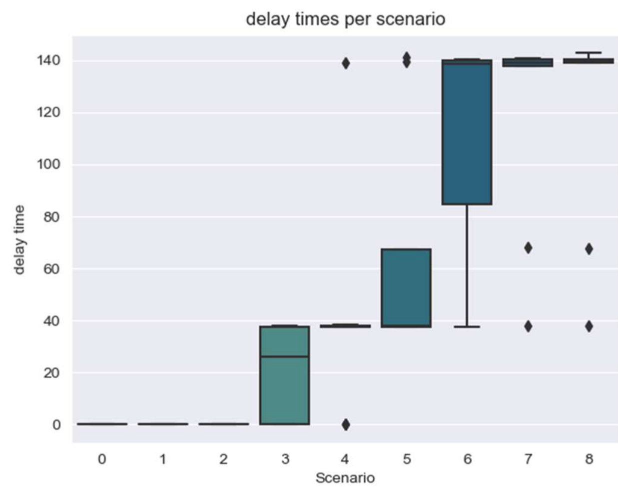

Figure 4: Boxplots of the average delay time for each scenario.

References

- Rhino Car Hire. (2023). Drive Left or Right - Which Countries Drive On the Left or Right. https://www.rhinocarhire.com/Drive-Smart-Blog/Drive-Left-or-Right.aspx
- Keller, N., & Hu, X. (2019). Towards data-driven simulation modelling for mobile agent-based systems. *ACM Transactions on Modeling and Computer Simulation* (TOMACS), 29(1), 1-26

Appendix

Variable	Metrics	Scenario 0	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8
driving time	min	585	585	585	585	585	637,595437	809,887843	848,014151	1154,15372
	max	585	585	585	733,744402	813,125196	930,495618	1151,04549	1431,94905	1985,09143
	mean	585	585	585	621,969979	655,379076	759,633848	991,761634	1091,87464	1591,99748
	std	0	0	0	47,4899565	69,2505067	98,3435858	124,343007	162,015094	234,393597
delay time	min	0	0	0	0	0	37,4125948	37,7113732	37,7992615	38,033402
	max	0	0	0	38,0833979	139,137207	141,343968	140,522161	140,895706	142,708522
	mean	0	0	0	20,2147403	40,3468894	64,1644019	111,939268	122,204826	122,847959
	std	0	0	0	18,6741287	38,0971961	41,884403	45,1598453	37,2268437	37,5512262

Table 3: Key metrics for overall driving and delay time in each Scenario. Scenario 0 relates to base-line scenario.

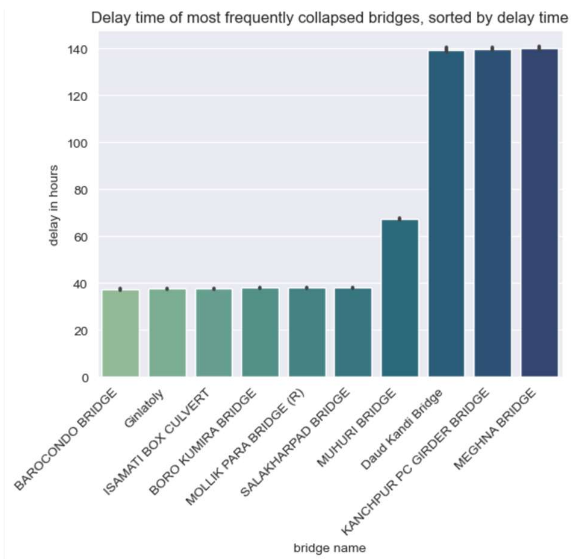

Figure 5: Delay times, caused by the 10 most frequently collapsed bridges.

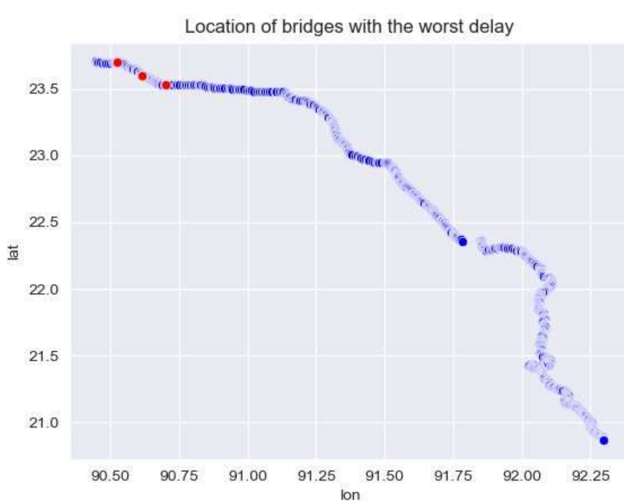

Figure 6: Location of the three worst bridges.