Travel Time and Fuel Consumption Optimization in Vehicle Routing under Fuzzy Congestion

A study of dynamic traffic routing with congestion levels calculated by applying fuzzy logic

Master Thesis

Helge McGough

Aalborg University
Mathematics-Economics

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Travel Time and Fuel Consumption Optimization in Vehicle Routing under Fuzzy Congestion

Theme:
A study of dynamic traffic routing with congestion levels calculated by applying fuzzy logic

Project Period:
Spring semester 2023

Project Group:
4.106 A3

Participants:
Helge McGough

Supervisor:
Inkyung Sung

Numbered Pages: 30

Date of Completion:
June 2, 2023

Synopsis:

In research, there is increasing interests in both efficient vehicle routing and quantified definitions of traffic congestion. As congestion can have a significant impact on most common routing parameters, especially travel time and fuel consumption, it could be beneficial to incorporate the traffic congestion into vehicle routing. As such, we define a fuzzy inference system to determine the level of congestion from the average speed and traffic density on a given road segment. The traffic congestion is then used to define penalties applied to the objective parameter in a Dijkstra's algorithm. The objective parameter to be minimized will be either fuel consumption or travel time, both described by a function of speed. The use of weighted moving average to forecast the input values of the routing is also investigated. The tests indicates that utilizing fuzzy determined traffic congestion can indeed improve the accuracy of routing algorithms, though when it comes to forecasting, there might be better options than the weighted moving average.

Preface

This project is written as a master thesis for a Mathematics-Economy student at Aalborg University. It is written in the spring semester of 2023.

The content of this report is freely available, but publication (with reference) may only be pursued due to agreement with the authors.

References are given using the Harvard system. All figures and tables are devised using `RStudio` or LaTeX for the purposes of this project.

All programming done is this project is done using R, and can be found in `https://github.com/hmcgou18/Travel-Time-and-Fuel-Consumption-Optimization-in-Vehicle-Routing-under-Fuzzy-Congestion.git`. Experiments are run on an Intel(R) Core(TM) i5-7200U CPU with 8GB of RAM.

Aalborg University, June 2, 2023.

Contents

1 | Introduction

Traffic congestion is a condition occurring in traffic streams characterized by slower speeds, delays in trip times, and restrained mobility, and is a pervasive issue, especially in larger cities. While there are several procedures to cope with traffic congestion, like building or extending traffic infrastructure, these procedures mostly just manage to slow the rate of which congestion intensifies, making traffic congestion inevitable [Kalinic and Krisp, 2019]. As such, it is reasonable to believe that traffic congestion can have a significant impact on the actual efficiency of a pre-planned route.

Vehicle routing is a vastly explored field of research, due to the immense presence of transportation in modern society. Be it a larger corporation needing to move huge loads of good between facilities or a single consumer needing to go from point A to B, an efficiently planned route will be desired. A common measure for the efficiency of routing methods is distance, due to lower distance leading to lower travel times. Minimizing fuel consumption has also been of very high interest in vehicle routing problems lately [Bektas and Laporte, 2011]. This is because a lower carbon emission lessens fuel costs and lessens the effect of climate change, meaning that making routes with low fuel consumption is a boon from both an economical and an environmental viewpoint. These measures, and other possible routing parameters, can be heavily affected by traffic congestion, and thus it would make sense to consider traffic congestion in vehicle routing.

While many ways of quantifying traffic congestion have been proposed, there is no single universally accepted definition [Aftabuzzaman, 2007]. Since congestion doesn't have a single uniform definition, it wouldn't seem logical to measure it uniformly. It should also be considered, that many traffic decisions are done under a great deal of imprecision and uncertainty. As such, utilizing a fuzzy inference system to combine different congestion measurements has been suggested, as this could account for said uncertainty in the individual congestion definitions [Berrouk et al., 2020].

As traffic conditions can be very dependent on the time of day, assuming a constant level of congestion during the travel might not be too good an idea. As such, incorporating time dynamics when determining the traffic congestion during vehicle routing might improve the efficiency of the resulting route.

1.1 Problem statement

Can the implementation of traffic congestion, determined by fuzzy inference system, improve the quality of vehicle routing? Furthermore, can addressing the changing time dynamics in traffic conditions, trhough the use of forecast data, further improve the quality of the vehicle routing?

1.2 literature review

1.2.1 Vehicle routing

As far as we are aware, a routing problem concerned with minimizing an objective function associated with travel speed was first outlined by [Fagerholt, 2001], who presents a schedule optimization algorithm for solving a shortest path problem on a sailing network, providing a fixed shipping route that minimizes the costs associated with sailing speed. Continuing his work, [Fagerholt et al., 2010] later formulate the optimization problem as a non-linear continuous program, and solve it using a non-linear programming solver and a shortest path algorithm. [Norstad et al., 2011] present a multi-start local search heuristic for solving a tramp ship routing and scheduling problem, wherein the subproblem of speed optimization along a fixed single ship route is solved by a recursive smoothing algorithm. This recursive smoothing algorithm is generalized in [Kramer et al., 2015b], by considering drivers wages, possible waiting times, and a non-fixed arrival time at the last customer. They further extends this algorithm in [Kramer et al., 2015a] to consider the departure times as decision variables.

In [Wen et al., 2016] they consider full-load tramp ship routing with variable speeds using a heuristic branch-and-price algorithm, and in [Fan et al., 2019] a tramp ship routing and speed optimization problem is investigated, using a Variable Neighborhood Genetic Simulated Annealing algorithm. The pollution-routing problem defined by [Bektas and Laporte, 2011] is the main application of these studies.

[Sung and Nielsen, 2019] considers time-dependent traffic conditions, based on traffic data from UK, with the possibility of changing the scheduled route mid-travel if required. An exact approach with approximation schemes is proposed as solution. In [Xiao and Konak, 2016] they study a green Vehicle Routing Problem with time-dependent traffic congestion, based on general congestion patterns. They attempt to minimize carbon emission by designing a MIP model that allows vehicles to stop at any point, so as to "wait out" the congestion and avoid excessive carbon emission from repeated acceleration. A hybrid algorithm of MIP and iterated neighborhood search is proposed to solve this problem.

1.2.2 Effect of congestion

The relationship between traffic congestion and routing parameters have also been a field of interest in research. [Errampalli et al., 2015] study the increase in fuel emission and travel time costs caused by a driving through a congested traffic state for various types of vehicles on varying multi-lane highways in India. In this study, congestion cost relationships have been developed between Congestion Factor and Volume-Capacity Ratio through the collection of exhaustive time related and fuel related data. [Treiber et al., 2007] investigate the difference in fuel consumption and travel time, on average, in free, bound, and congested traffic states, using NGSIM trajectory data. The fuel consumption is determined by a function of instantaneous velocity and acceleration. Curiously, their results found that the bound state had a lower fuel consumption on average than the free flow state.

[Jereb et al., 2017] look at the impact of traffic flow on fuel emission in urban environments, by studying the difference in traffic flow between "green light" and "red light" phases at an intersection with traffic lights. [McKnight et al., 2003] attempt to quantify the additional travel time that buses might need due to traffic congestion, by developing a regression model that estimates the travel time rate (in minutes per mile) of a bus as a function of car traffic rate, number of passengers boarding per mile, and the number of bus stops per mile.

1.2.3 Definitions for traffic congestion

There have been many studies proposing different ways of defining traffic congestion. Some of them can be seen in [Aftabuzzaman, 2007], who have reviewed and categorized congestion definitions into 3 broad categories: Demand-Capacity related, where a congested state is when the capacity of the road is exceeded, Delay-Travel time related, where the difference in actual travel time from an ideal free flow state are used to determine level of congestion, and Cost related, where traffic congestion refers to the incremental costs resulting from interference among road users. [Bovy and Salomon, 2002] defines congestion as a state of traffic flow, characterized by low speeds and high densities.

Lately, fuzzy logic have been implemented to address the uncertainties involved in quantifying traffic congestion. In particular, the use of fuzzy inference systems in determining congestion have been investigated by [Toan and Wong, 2021], [Kalinic and Krisp, 2019], and [Berrouk et al., 2020]. [Toan and Wong, 2021] uses speed and density as input variables with a crisp congestion index in $[0, 1]$ as output. All fuzzy parameters are partitioned into 5 categories through membership functions. [Kalinic and Krisp, 2019] uses traffic flux and density as input values to determine a fuzzy congestion, while having 7 partitions for each parameter. The fuzzy systems in both these studies are tested using inductive loop detector data collected by Mobile Century on Interstate 880, California. In [Berrouk et al., 2020] they have three input variables: volume-to-capacity ratio, decreased speed ratio, and speed ratio, all with varying number of fuzzy partitions. Their fuzzy system is tested on more urban traffic data from Austin, Texas. [Tišljarić et al., 2021] determines traffic congestion based on the center of mass in a speed transition matrix using a fuzzy inference system. The input values are the x and y coordinates of the center of mass in a speed transition matrix.

1.3 Data acquisition and preparation

To test the proposed method and fuzzy system, we will use For this project, we use data from the "Radar traffic Counts" set from Austin, Texas, which is freely provided by the Austin Department of Transportation at `https://data.austintexas.gov/Transportation-and-Mobility/Radar-Traffic-Counts/i626-g7ub`.

This data contains data for volume and average speed of vehicles, measured in 15 minute intervals. The data is collected by Wavetronix radar sensors located at 18 different locations. This data have been collected from November 2017 to September 2021, resulting in a rather large data set of 1.452.901 rows of data, after aggregation. As the data is collected at each location on an "as needed" frequency, there are time points that only contains data for few of the road segments. As such we limit our testing to the year 2018, as this year to include the most road segments at each collection period. This give us 457,518 rows of data, spread across 30,043 time points, to make traffic networks from, with an average of $15 - 16$ nodes per network.

2 | Model description

In this project we will focus on road traffic networks, meaning that constructed solutions will be considering land vehicles only. The network is presented as a set of nodes and a set of arcs connecting said nodes. The vehicles in the network will only be able to travel trough these arcs. For a given vehicle there will be both a start and a destination node. The goal of the driver is assumed to simply be arriving at their destination in the least costly way possible. As such all nodes, besides the start and destination, in the network will be optional intermediate nodes, none of which are specifically required to be visited, but only exist to link the start and end node together. We also assume that there will be no idle time at these intermediate nodes. The vehicle speed will be assumed constant during each arc, and can only be changed at intermediate nodes. Rerouting will not be considered possible during the scope of this project, and thus the route must be determined during the planning phase.

2.1 Optimization problem

Let us have $n_i \in N$ denote the set of nodes in a given traffic network. $(i, j) \in A$ will then denote the set of arcs connecting the nodes in N. Specifically, arc (i, j) will denote the road segment going from node n_i towards node n_j, and will thus have the traffic conditions and road length associated with node n_j. The start and destination node will be denoted by $n_{start}, n_{end} \in N$ respectively. A sequence of arcs connecting n_{start} and n_{end}, which shall be called a route, will be denoted as A', where $\Big((N_{start}, n_i), \ldots, (n_j, n_{end})\Big) = A' \subset A$. The nodes contained in the route A' will likewise be denoted as N', where $n_{start}, \ldots, n_{end} \in N' \subset N$. t_i will denote the time point at which the vehicle leaves node n_i, and $t_i \in T', n_i \in N'$ will be the set of leaving times in the associated route A'.

The goal of the vehicle routing will be the minimization of either the travel time or the fuel consumption. The length, in kilometer, of arc (i, j) will be denoted $l_{(i,j)}$. Note that since arc (i, j) represents a road segment associated with node n_j, we have that $l_{(i,j)} = l_{(k,j)}, \forall n_i, n_k, n_j \in N \wedge \forall (i, j), (k, j) \in A$. As such, for an arc (i, j), the travel time over length $l_{(i,j)}$ with constant speed v will be denoted as $s^v_{l_{(i,j)}}$, while $c^v_{l_{(i,j)}}$ will be fuel emission over the $l_{(i,j)}$ kilometers at speed v. The traffic congestion on arc (i, j) at time t_i will be denoted as $f^{t_i}_{(i,j)}$. The impact of congestion on the fuel consumption and travel time will be represented by penalty factors, which are determined based on the traffic congestion. These penalty factors are denoted by $p^c_{f^{t_i}_{i,j}}$ and $p^s_{f^{t_i}_{i,j}}$ for fuel consumption and travel time, respectively. These penalties will be further defined in subsection 2.1.1. Using these notations, we define an objective function for minimizing travel time as,

$$O_s = min\left(\sum_{(i,j)\in A', t_i \in T'} s^v_{l_{i,j}} \cdot p^s_{f^{t_i}_{i,j}}\right), \tag{2.1}$$

where $s^v_{l_{i,j}}$, measured in seconds, is determined by

$$s^v_{l_{i,j}} = \frac{l_{(i,j)}}{v}. \tag{2.2}$$

It is reasonable to assume that travel time is affected by traffic congestion, as slower speeds are a common characteristic of congested state, and vehicle is the parameter from which travel time is determined. This speed reduction should be captured by the speed data collected by the traffic sensors. However, as it wouldn't be too strange for the collected data to contain imprecisions, since installing sensors in optimum locations is still quite a challenge [Gentili and Mirchandani, 2012], [Hu et al., 2009], especially during congested states, where the traffic density is high. As such, we hope that penalizing the travel time, will increase the efficiency of a route determined through travel time.

We also define an objective function for minimization of fuel consumption,

$$O_c = min\left(\sum_{(i,j)\in A', t_i \in T'} c^v_{l_{i,j}} \cdot p^c_{f^{t_i}_{i,j}}\right) \tag{2.3}$$

where the fuel consumption for road length $l_{(i,j)}$, measured in grams, is determined trough the following function developed for a 9.58-ton diesel truck through real-life empirical studies [Coyle, 2007]

$$c^v_{l_{i,j}} = (871 - 16 \cdot v + 0.143 \cdot v^2 + 32031 \cdot v^{(-2)} + 1.79961 \cdot 10^{(-2)} \cdot 9.58) \cdot l_{(i,j)} \tag{2.4}$$

This function is based on the emission model developed by [Hickman et al., 1999], and has also been applied by [Xiao and Konak, 2016]. This function assumes that average vehicle speed is enough to determine the fuel consumption of an arc (i, j). This can of course be considered a rather "rough" assumption to make, as fuel consumption is heavily affected by other factors as well, such as braking and accelerating frequency. However, as we don't have access a function taking acceleration into account, nor any such data, we instead try to penalise the speed based fuel consumption function according to congestion level. Using (2.4) to determine fuel consumption also requires the assumption that the vehicles will be 9.58-ton diesel trucks. This assumption can be changed based on the fuel consumption function used, as the rest of the model doesn't assume a specific fuel function.

The objective functions (2.1) and (2.3) is minimized while subjected to the following constraints:

- $t_{i_{start}} = 0$, meaning that we begin the route at time 0.

- $\sum_{(i,j)} 1_{A'}(i,j) = 0 \vee 1, \forall (i,j) \in A$, meaning that we don't have to traverse every arc in the network A during our route A', and that each arc can only be visited once at most.
- $\sum_{n_i} 1_{N'}(i) = 0 \vee 1, \forall i \in N$, like above this means that we won't have to visit every node in the network N during our route N', and that each node can only be visited once at most.
- $\sum_{(i,j)} 1_{A'}\left((i,j) \wedge (j,i)\right) = 0, \forall (i,j) \in A$, secures that we won't be traversing the other way of a road segment that we have already traversed.

2.1.1 Defining the penalties

The penalties $p^s_{f^{t_i}_{i,j}}$ and $p^c_{f^{t_i}_{i,j}}$ are factors meant to represent the increase in travel time and fuel consumption, respectively, caused by the level of congestion on the given road segment. As traffic congestion results in worse travel times and fuel emissions, it makes sense for the penalties to increase along with the congestion index. Of course, we must still define how the travel time and fuel emission increases along a growing congestion index. Multiplying the congestion index directly on the objective parameter would be a simple yet naive solution, as it assumes a linear relationship between the penalties and the traffic congestion.

It is more reasonable to assume, that the growth in penalties follows a more exponentially increasing relationship with the traffic congestion. The growth in travel time and fuel consumption along increasing levels of congestion hasn't been sufficiently researched, although there are studies on the difference in travel time and fuel consumption between free flow and congested traffic state [Errampalli et al., 2015]. As such we will need to approximate how the travel times and fuel emission increases along the mid-range levels of traffic congestion.

We have first approximated the relationship between the travel time and congestion index, using the Austin data. This is done by calculating the difference in travel time when driving at the average speed, taken from the data, and the maximum allowed speed, which is set to be the legal speed limit of the associated road segment. Thus we assume that a driver will aim to drive as fast as possible, thus minimizing travel time, while still upholding the speed limit. While we believe that it is reasonable to assume that most drivers will prioritize minimizing travel time, and thus maximising speed, we are aware that many have no problem driving above the speed limit, but let's keep things legal. The differences in travel time to congestion level can be seen in Figure 2.1.

On average we see the following, Table 2.1, increase in travel time at the given levels of congestion.

The travel time penalty $p^s_{f^{t_i}_{i,j}}$ will then be determined based on the fuzzy membership of the congestion level to each of the congestion levels from Table 2.1. The

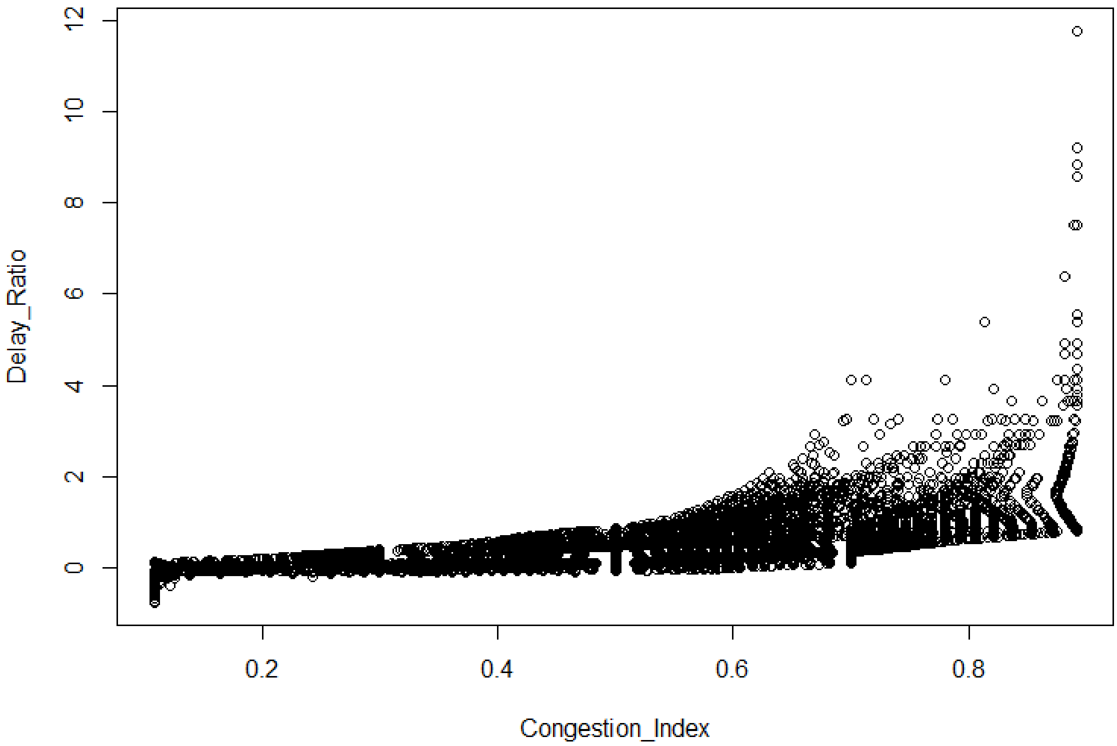

Figure 2.1: Delay ratio to congestion index.

Congestion	delay in travel time
very_light	3%
light	11%
moderate	22%
heavy	45%
very_heavy	79%

Table 2.1: Average delay in percentage to each linguistic partition of the congestion index.

fuzzy membership function and the predicates are described in chapter 3. The travel time penalty can now be calculated by the following equation:

$$p^s_{f^{t_i}_{i,j}} = 1 + (0.03 \cdot m_{very\_light}(f^{t_i}_{i,j}) + 0.11 \cdot m_{light}(f^{t_i}_{i,j}) + 0.22 \cdot m_{moderate}(f^{t_i}_{i,j}) + 0.45 \cdot m_{heavy}(f^{t_i}_{i,j}) + 0.79 \cdot m_{very\_heavy}(f^{t_i}_{i,j})) \tag{2.5}$$

where $m_{very\_light}(f^{t_i}_{i,j})$ is the degree of membership, represented by a value in $[0, 1]$, for the congestion index $f^{t_i}_{i,j}$ in belonging to the very_light congestion fuzzy set. The way these membership functions are defined, ensures that the total degree of membership will be equal to 1.

As for determining the scaling of the fuel emissions penalty $p^c_{f^{t_i}_{i,j}}$, it gets a bit more tricky. Our previous assumption that the typical driver focus on minimizing travel time, means that we cannot make the same assumption regarding fuel emission.

We also find it more reasonable to assume that a driver would optimize travel time over fuel emission, as unlike travel time, which decreases as speed increases, the fuel emission have a parabolic relationship with speed. This means that there is a specific speed that results in the least fuel emission, and as such increasing the speed won't necessarily decrease fuel emission. Thus the optimal speed for minimizing fuel consumption is heavily dependent on the specific vehicle, and I doubt that the average driver is aware of their vehicles optimal speed for fuel reduction. All this aside, we don't have any data directly related to the fuel consumption. As such we can't directly use the Austin data to approximate the penalty for fuel emission.

[Errampalli et al., 2015] have studied the difference in fuel emission between very_light and very_heavy congestion in traffic, for different types of vehicles. Specifically they observed a 120.7% increase in fuel consumption for heavy commercial vehicles, which are the type of vehicles used in the fuel emission function (2.4). As we don't have a way of knowing the specific scale differences between the fuel emission penalty at very_heavy congestion and the other levels, they will be assumed to be of the same scale as in the travel time penalty. This gives us the following fuel emission penalty function

$$\begin{aligned} p^{c}_{f^{t_i}_{i,j}} =& 1 + ((1.207 \cdot (3/79)) \cdot m_{very\_light}(f^{t_i}_{i,j}) + (1.207 \cdot (11/79)) \cdot m_{light}(f^{t_i}_{i,j}) \\ &+ (1.207 \cdot (22/79)) \cdot m_{moderate}(f^{t_i}_{i,j}) + (1.207 \cdot (45/79)) \cdot m_{heavy}(f^{t_i}_{i,j}) \\ &+ 1.207 \cdot m_{very\_heavy}(f^{t_i}_{i,j})) \end{aligned} \tag{2.6}$$

where the terminology is the same as in (2.5). Again the way our membership functions are defined, ensures that the total degree of membership will equal 1. Note that changing the fuel consumption function (2.4) to fit another vehicle type would require changing the factor of fuel increase in very_heavy congestion. The value of this factor at other vehicle types can be seen in [Errampalli et al., 2015].

3 | Fuzzy inference system for detecting traffic congestion levels

In this chapter, we will present the applied fuzzy logic system that is used to determine the congestion levels of the road segments.

The fuzzy system used in this project will be of the Mamdani type [Mamdani, 1974], which is the most commonly used fuzzy control system. A Mamdani fuzzy system roughly follows these three steps:

1. Fuzzify input values into membership functions.

2. Use IF-THEN rules to compute fuzzy output functions.

3. Defuzzify the fuzzy output into a "crisp" output value

3.1 Input and output parameters

Input parameters are crisp values that can be used to describe the output parameter. In this case I want my input parameters to be quantitative real data that can help in defining traffic congestion.

Speed is a clear indicator for traffic congestion, as reduction in vehicle speed is basically the trademark of traffic congestion. In fact, congestion can outright be described as a function of speed reduction [Council, 2000], making average vehicle speed an obvious choice as the first input parameter in our fuzzy system. Many studies utilizes speed in determining traffic congestion, ([Toan and Wong, 2021], [Kukadapwar and Parbat, 2015], [Bovy and Salomon, 2002], [Rao and Rao, 2012], [Hamad and Kiku
Note that the speed from the given data, which is measured in miles per hour, will be converted to kilometer per hours.

Another indicator on the traffic could be the traffic flux, which have been used in the fuzzy inference systems of previous studies ([Kalinic and Krisp, 2019], [Rao and Rao, 2012], [Berrouk et al., 2020]). However, the relationship between traffic flux and congestion is a little more complex, as a low flux can indicate both a congested state or a free flow state. As such the measured traffic flux won't properly reflect the traffic demand under congested state, the estimated demand can become misleading. The parabolic relationship between speed and flux also make it more difficult to formulate fuzzy rules using speed and flux, as whether a low flux indicates a low or high congestion level will depend on the vehicle speed, making the fuzzy system more dependent on the speed level. Hence, we will not use the traffic flux as input parameter in our fuzzy inference system.

Another viable option for an input parameter would be the traffic density. The relationship between speed and density is monotonic with a negative slope, making it easier to formulate fuzzy rules using these two as input parameters. Traffic density is also a common choice for evaluation of traffic congestion ([Toan and Wong, 2021], [Arabani and Pourzeynali, 2005], [Bovy and Salomon, 2002]), as it directly affects maneuverability in the traffic stream. While density is not directly given from the Austin data, it can be derived from the flux and speed data by the following equation

$$k = \frac{q}{v}, \tag{3.1}$$

where k is the density measured in vehicles per kilometer, v is average vehicle speed, and q is traffic flux. Thus, density will be the second input parameter of our fuzzy inference system.

The fuzzy inference system also requires an output parameter to be determined based on the membership values of the input parameters. This output parameter shall be a congestion index, expressed as a crisp value in $[0, 1]$, where 0 means completely free flow and 1 means complete congestion.

3.2 Membership functions

Both input parameters and the output parameter will now be assigned 5 linguistic predicates each, which are

$$\begin{aligned} Speed &= \{very\_slow, slow, medium, fast, very\_fast\} \\ Density &= \{very\_low, low, medium, high, very\_high\} \\ Congestion &= \{very\_light, light, moderate, heavy, very\_heavy\} \end{aligned}$$

Each of these linguistic predicates are characterized through a membership function which calculates the degree of membership, represented by a value in $[0, 1]$, of which a given crisp value belongs to the predicate. The membership functions in our fuzzy inference system are made in reference to the ones from [Toan and Wong, 2021].

As can be seen in Figure 3.1, the universe of discourse for density is represented by three triangular membership functions for the middle ranges (light, medium and high), and two half-triangular functions for the boundary ranges (very_light and very_heavy). The membership functions for speed and congestion index will have similar structures.

What constitutes an e.g. "fast" vehicle speed is dependent on the speed limit. As our fuzzy system will be used in traffic networks containing road segments with varying speed limits, we will need a set of speed membership functions for each speed limit in our networks. All these membership functions are shown in Figure 3.2.

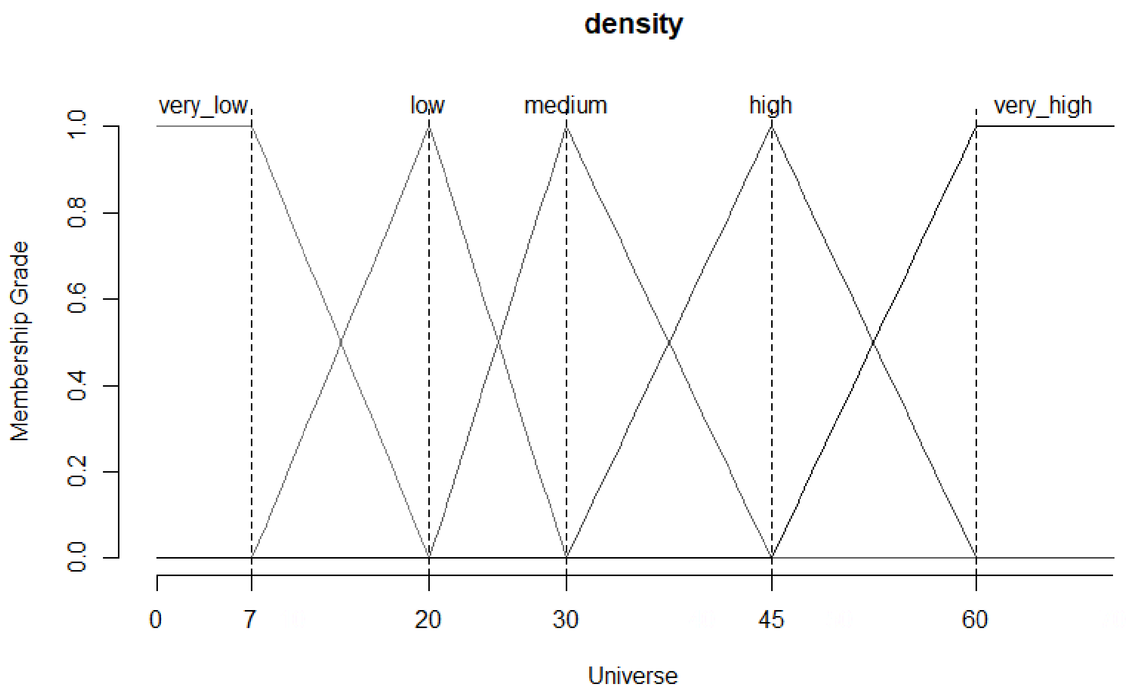

Figure 3.1: The fuzzy partition for density

The membership functions for the congestion index is shown in Figure 3.3. The conditions that each congestion predicate represent is described in the following list:

- "Very_light" congestion refers to free or near-free flow traffic state, where there is almost restrictions on the vehicles.
- "Light" congestion is when when the speed starts to reduce and the freedom to maneuver within the traffic stream is becoming noticeably limited.
- "Moderate" congestion describes conditions where speed deceases significantly, density increases quickly with increasing flows, and maneuverability within the traffic stream is limited.
- "Heavy" congestion indicates breakdowns in vehicular flow at which point queues may form with potential propagation upstream.
- "Very_heavy" congestion represents an extreme breakdown of flow and very low traffic dynamics.

It should be noted that the terms are imprecisely defined following the fuzzy logic concept, and there are no clear cuts between the fuzzy congestion levels. Degree of membership represents a truth value that is primarily governed by subjective degree of belief. In this regard, the overlaps between the fuzzy sets are designed such that the sum of membership grades for adjacent fuzzy sets at any point in the overlapping sections equals 1. It follows that the degree of membership in a particular fuzzy set approaches 0 as that of the adjacent fuzzy set approaches 1.

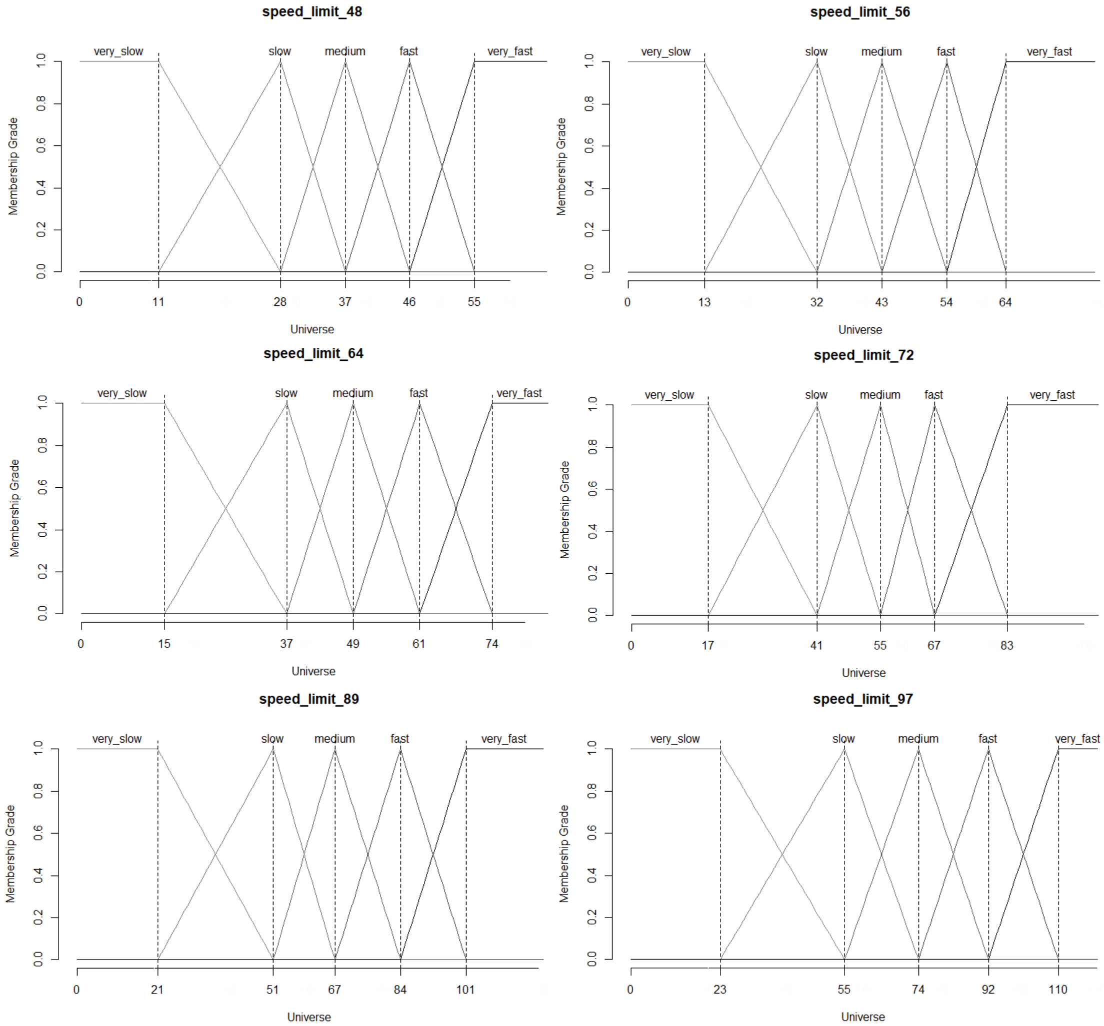

Figure 3.2: Fuzzy partition for speed, at each of the speed limits for the road segments

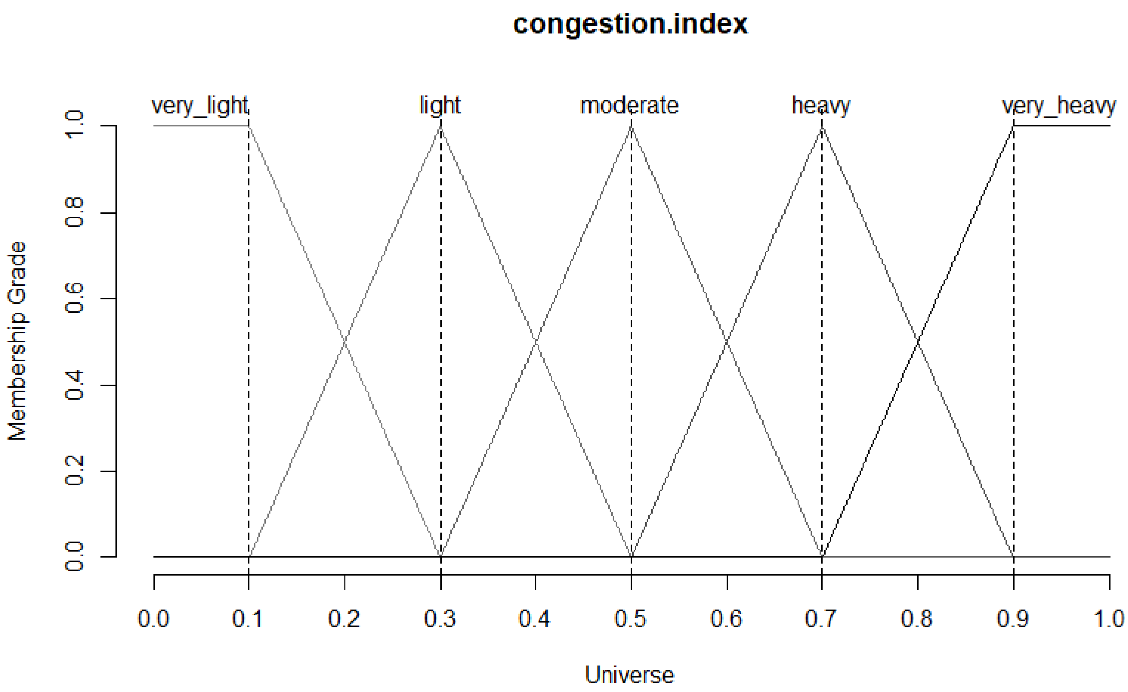

Figure 3.3: Fuzzy partition for congestion index.

3.3 The fuzzy rules

The purpose of fuzzy rules is to map the fuzzyfied input values to the desired fuzzy output predicate. Using IF-THEN rules, the two fuzzyfied inputs will be combined via the AND operator to determine a fuzzy output. Being connected with AND operator, the membership values of the rules are calculated using MIN operation in fuzzy logic. Each rule combination is evaluated parallel, applying a minimum implication operator, giving us an output fuzzy set for each rule combination. These different output fuzzy sets are then combined into a single fuzzy set via maximum aggregation. Lastly this aggregated fuzzy set is defuzzyfied into a crisp congestion index using the centroid method, which returns the center of the aggregated area under the curve as our crisp output. The fuzzy rules are summarized in Table 3.1.

Density / Speed	very_slow	slow	medium	fast	very_fast
very_high	very_heavy	very_heavy	heavy	heavy	moderate
high	very_heavy	heavy	heavy	moderate	moderate
medium	heavy	moderate	moderate	moderate	light
low	very_heavy	moderate	light	light	very_light
very_low	very_heavy	very_heavy	light	very_light	very_light

Table 3.1: Fuzzy IF-THEN rules in the form of a decision matrix.

Most of these fuzzy rules are also used by [Toan and Wong, 2021]. However, they did not include the rule combinations: (very_fast speed - very_high density), (very_fast speed - high density), (fast speed - very_high density). These were not included because they were considered to be too unlikely to occur in real life traffic. These combinations have, however, been included in our fuzzy system, and their corresponding congestion predicate have been evaluated to the best of our abilities. The rule combinations of (very_slow speed - very_low density), (low speed - very_low density), and (very_slow speed - low density), are also unlikely to occur naturally, but may happen under special circumstances, like incidents or roadwork. These three rule combinations will thus correspond to a very_heavy congestion level, as the circumstances under which they can happen would most likely slow down traffic to heavily congested levels, if not make the road entirely untraversable.

4 | Vehicle routing

In this chapter, the process of finding the best path by the objective function is presented.

4.1 Traffic networks

The proposed method will be tested on traffic networks, randomly generated from the data available at a given time point. The nodes in these networks will be the sensors from the data, labelled as the associated "KITS.ID". Any arcs going towards a particular node will be considered a road segment associated with that nodes "KITS.ID". The traffic networks will be generated using the "random.graph()" function from the "bnlearn" package in R programming. For each pair of nodes there will be a given probability of an arc appearing to connect said pair of nodes. An example of a randomly generated road network is shown in Figure 4.1. This network is made with a given arc probability of 30% for all pairs. Note that all arcs are undirected.

4.2 Finding best route

The best route will be found by a shortest path algorithm, specifically the well known Dijkstra's algorithm. A starting node and a destination node will need to be specified. During the initial testing, we will not account for change in traffic condition over time. As we are using relying the Dijkstra's algorithm for the routing, our proposed method will have a time complexity of $\mathcal{O}(n)^2$. n is the amount of nodes in the traffic network, which won't be exceeding 18 during our testing.

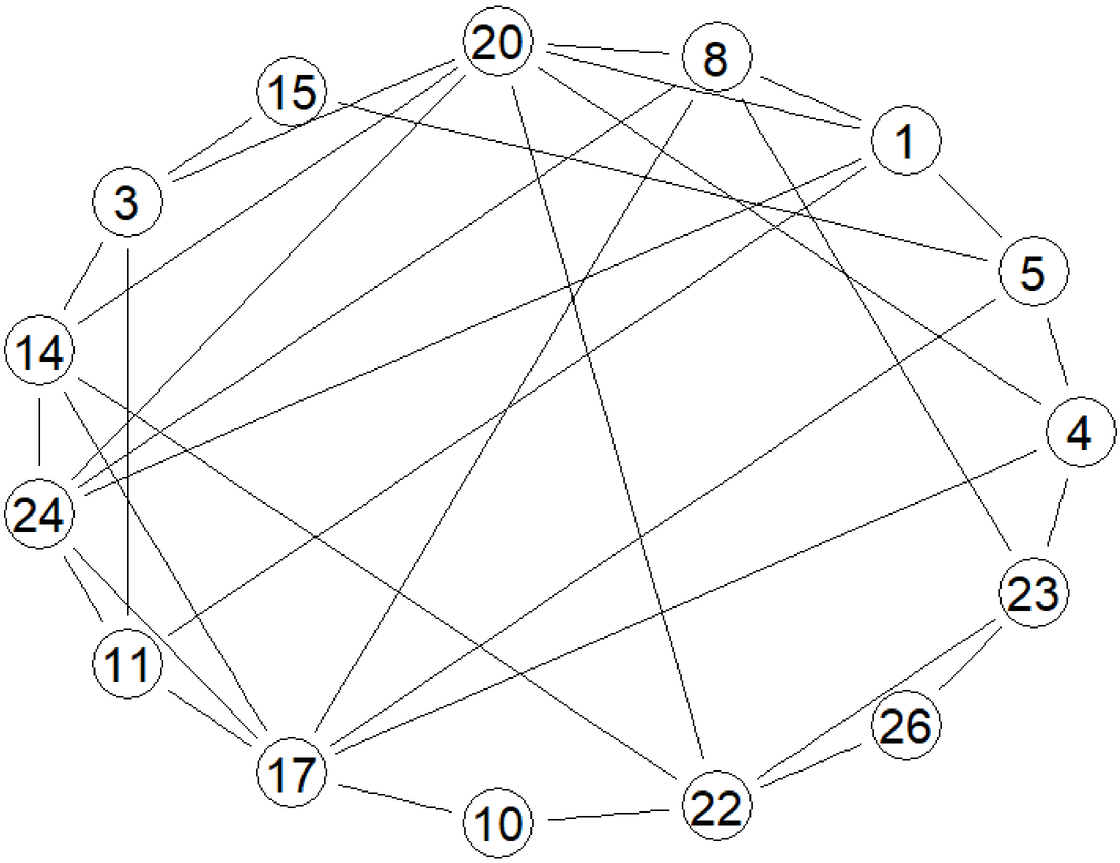

Figure 4.1: An example of a random traffic network on the standard setting.

4.3 Comparison

In order to have something to measure the proposed measure against, we create a benchmark. In order to measure the improvement caused by using congestion in vehicle routing, we will be comparing the objective function with a benchmark. This benchmark is also found by Dijkstra's algorithm, but without considering the traffic congestion, meaning that the travel times and fuel consumption determined from the data won't be penalized. As such the benchmark chooses its routes based on (4.1) or (4.2).

$$B_c = min\left(\sum_{(i,j)\in A} c^v_{i,j}\right) \tag{4.1}$$

$$B_s = min\left(\sum_{(i,j)\in A} s^v_{i,j}\right). \tag{4.2}$$

4.4 Target measure

To evaluate the accuracy of the proposed method and the benchmark we will need a "goal", ideally being the completely accurate optimal routes and objective values. We can then measure the accuracy of the proposed method and the benchmark by their proximity to said "goal". To get such a goal, we use Greenshield's model, (4.3), to determine the "accurate" vehicle speeds, based on the density data [Greenshields et al., 1935].

$$v = v_f \cdot \left(1 - \frac{k}{k_j}\right) \tag{4.3}$$

where k is density determined from the data, k_j is the jam density, set to $k_j = 150$ based on the default value from Wu's Fundamental Diagram, [Wu, 2002], and v_f is the free flow speed, set to be the speed limit of the relevant road segment. The speeds determined by (4.3) will, for lack of better options, be considered the "accurate" speed data, and are thus used to determine the "accurate" best routes in the traffic networks that will serve as our "goal" for the proposed method and the benchmark.

4.5 Forecasting

As traffic conditions aren't static, it would make sense to take potential changes to said conditions into considerations during the route planning. As such we will also implement a weighted moving average forecasting. We forecast the traffic congestion in 15 minute intervals, as that is the intervals in which the data is updated. Aside from the congestion index, we will also forecast the speed of the relevant road segment, for use in the objective function. The weighted moving average will be based on the 3 previous values with weight 3 for the immediately past value $t-1$, weight 2 for $t-2$, and weight 1 for $t-3$. This is shown in the following equation

$$\begin{aligned}
v_t &= \frac{v_{t-1} * 3 + v_{t-2} * 2 + v_{t-3} * 1}{3+2+1} \\
k_t &= \frac{k_{t-1} * 3 + k_{t-2} * 2 + k_{t-3} * 1}{3+2+1} \\
f_t &= \frac{f_{t-1} * 3 + f_{t-2} * 2 + k_{t-3} * 1}{3+2+1},
\end{aligned}$$

where v_t, k_t and f_t is the forecast speed, density and congestion to time t, respectively. Using these forecast parameters, we will also determine a forecast version of the proposed method to compare with the non-forecast version. The forecast version will also be compared to forecast versions of the benchmark and "goal" described above.

5 | Results

We now investigate the performance of the proposed method against the benchmark, both with and without forecasting

For our tests we have defined a standard setting for the traffic networks on which we will investigate the performance of our proposed method. In the standard setting the networks are generated with 30% arc probability. For each iteration, the starting time, as well as the start and destination nodes, will be chosen at random from the available data. The lengths of each sensors road segment is randomly sampled from the discrete interval $[5, 10]$.

All the results are averages over 500 iterations of the relevant test, and all test runs are seeded for better comparison between test settings.

The terminology used in the columns of the result tables are as follows:

- "forecast": A yes/no variable for whether or not forecasting was used to get the associated results.
- "method": The method used to get the associated results.
- "parameter": Whether the method is judged based on the fuel emission or travel time.
- "routes": The percentage of iterations where a different route was chosen between the given method and the target.
- "objective": The average percentagewise difference in the objective function value between the route chosen by the given method and the target route.
- "gap route": The difference in precision of finding the target route between the proposed method and the benchmark.
- "gap obj": The difference in precision of obtaining the same objective value as the target route between the proposed method and the benchmark.

forecast	method	parameter	routes	gap route	objective	gap obj
no	proposed	fuel emission	7.23%	1.65%	-2.30%	0.5%
no	benchmark	fuel emission	8.88%		-2.80%	
yes	proposed	fuel emission	7.59%	2.01	-2.38%	0.73%
yes	benchmark	fuel emission	9.60%		-3.11%	
no	proposed	travel time	14.26%	4.54%	-5.21%	-2.96%
no	benchmark	travel time	18.80%		-2.25%	
yes	proposed	travel time	14.06%	5.36%	-4.45%	-1.68%
yes	benchmark	travel time	19.42%		-2.77%	

Table 5.1: Results from the standard setting.

When minimizing fuel consumption, the proposed method chooses the same route as the target more often than the benchmark does, and the fuel consumption in the different routes found by the proposed method are on average closer to the fuel consumption of the target routes. This indicates that the proposed method leads to more accuracy in vehicle routing, although not by much.

When forecasting the values the precision gaps widens, so the proposed method is even more accurate compared to the benchmark, but also a little bit less accurate than the non-forecast proposed method.

When minimizing travel time the proposed method more often picks the same route as the target than the benchmark does. In cases when another route than the target is found, the travel time in the benchmark routes tend to be closer to the target value. As such, while the proposed method finds the same route as the target more often, when it doesn't it will on average lead to a larger delay in travel time compared to the benchmark.

While forecasting do give a slight improvement in the accuracy of the proposed method, the delay from a missed route is still higher than in the benchmark.

6 | Discussion

In this project, Greenshield's model, [Greenshields et al., 1935], was chosen as the target measure primarily because of a lack of better options. Greenshield's model determines the speed as a function of the traffic density, and as such it can be considered a speed determined by the congestion level, when congestion is measured as the traffic density only. Of course, this model probably isn't necessarily accurate itself, as it is also based on collected data, in which there can be imprecisions, just like we assume in our data.

Travel time was measured as a function of average speed, which we got from the data. While it is reasonable to assume that the measured average speed is already affected by congestion, it is also reasonable to assume that high traffic density can lead to imprecisions in the sensors measurment of said speed data. Especially as there were cases in the data where there was a high speed at high density, which should be very unlikely. As such we experimented with penalizing the travel time, to make up for said imprecision.

For measuring the fuel consumption, it would have been useful to incorporate e.g. acceleration and braking frequency, as an average speed is not enough to determine an accurate fuel consumption. We couldn't find an applicable method that included such parameters, and even if we could we still don't have access to any such data anyways. As such, we chose to use a fuel consumption function based on average speed only, as it was both available and easily applicable on the proposed method.

The penalty factors could have significant impact on the accuracy of the proposed method, since the penalty factors reflect the extra cost imposed on the parameters due to the congestion level. As such, having accurate penalty factors is of utmost importance in the proposed method, as incorporating congestion in vehicle routing wouldn't matter if the effect of said congestion in the model didn't properly reflect reality. While the penalty factors used in the proposed model were made to be as accurate as possible to best of our ability, there are still reasons to suspect imperfections in said penalty factors. The general effect of traffic congestion is likely to depend on the local area, as general driver behaviour can vary based on the local culture. While we tested using data from Austin, Texas, the penalty for the fuel consumption were partly based on data from India, which can have lowered the accuracy of said penalty. General traffic expertise of the decision maker would also help in determining accurate penalties. As such, While the penalty factors designed in this study is made to reflect the real impact of traffic congestion on the parameters, they could probably still be improved upon, by local traffic experts with access to proper local data.

The forecasting method chosen was a rather simple weighted moving average. While it wasn't too much, incorporating a weighted moving average in the proposed method still lead to a slight improvement in the chosen routes when minimizing travel time, and without really increasing the computation time. This wasn't much of an improvement however, and forecasting seemed to actually worsen the accuracy when minimizing fuel consumption. The weights of the utilized moving average were, however, chosen somewhat arbitrarily to put more impact on the more recent data, meaning that more properly tuning the weights might increase the effectiveness of the forecasting. It probably also wouldn't hurt to test out other short term forecasting methods, such as exponential smoothing. A complex forecast with a long computation time probably won't be worth it however, as the route shouldn't take too long to compute when working with dynamic time, which you most likely are if forecasting is involved. Investegating long term forecasting is most likely unnecessary, as incorporating traffic congestion in vehicle is more useful in urban environments, where the length of a routes travel time is limited.

The fuzzy sets and rules used in the inference system of this project is heavily inspired by the system used by [Toan and Wong, 2021], who made their system with a highway in mind instead of more urban networks, where traffic congestion would most likely matter more in routing. While we have scaled the fuzzy set for speed to fit the speed limits of the road segments in the data, it is possible that the fuzzy sets for density can also be different for each road segment. Defining density fuzzy sets for each road segment would however require some knowledge regarding the capacities of said road segments, which I did not have access to.

7 | Conclusion

This project was concerned with investigating if the implementation of traffic congestion, determined by fuzzy inference system, could improve the quality of vehicle routing. This have been investigated in routing problems focused on minimizing travel time and fuel consumption, separately. While we do find some improvement in utilizing traffic congestion in vehicle routing, it should be noted that the proposed model is limited by certain assumptions, that aren't entirely reasonable to make. Primarily it would be recommended to utilize a fuel consumption determined by more than just the average speed. However, such data may be difficult to access, and methods to determine fuel consumption based on acceleration or braking frequency may be more complex, than the function used in this project.

We were also concerned with investigating the use of forecast values in vehicle routing with traffic congestion, and while the weighted moving average used in the tests did not yield much results, there may still be potential in other short term forecasting methods.

While fuzzy inference systems are useful in determining the traffic congestion, the effect of different levels of congestion, as well as the fuzzy predicates and membership functions of the inputs, will need to be defined. As such there are plenty of decisions involved in defining a system for incorporating traffic congestion into vehicle routing. The correct decisions can be very dependent on the locality of the traffic network, and as such traffic expertise and general knowledge about the local area can be greatly beneficial in defining a fuzzy system for traffic congestion.

8 | Bibliography

[Aftabuzzaman, 2007] Aftabuzzaman, M. (2007). Measuring traffic congestion- a critical review. *Proceedings of the 30th Australasian Transport Research Forum (2007), p. 16.* https://www.researchgate.net/publication/282780229_Measuring_traffic_congestion-_A_critical_review.

[Arabani and Pourzeynali, 2005] Arabani, M. and Pourzeynali, S. (2005). Fuzzy logic methodology to evaluate the service level of freeways basic segments. *Iran. J. Sci. Technol. Trans. B, Eng., 29.* https://www.researchgate.net/publication/251742924_Fuzzy_logic_methodology_to_evaluate_the_service_level_of_freeways_basic_segments.

[Bektas and Laporte, 2011] Bektas, T. and Laporte, G. (2011). The pollution-routing problem. *Transportation Research Part B: Methodological, 45(8),1232-1250.* https://doi.org/10.1016/j.trb.2011.02.004.

[Berrouk et al., 2020] Berrouk, S., Fazziki, A., and Sadgal, M. (2020). Fuzzy-based approach for assessing traffic congestion in urban areas. *Springer Nature - PMC COVID-19 Collection.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7340959/.

[Bovy and Salomon, 2002] Bovy, P. and Salomon, I. (2002). Congestion in europe: measurements, patterns and policies. *In Monograph Travel Behaviour: spatial patterns, congestion and modelling. Pp.143-179.* http://resolver.tudelft.nl/uuid:ed2da712-8b04-4576-85c8-4fbe518c061d.

[Council, 2000] Council, N. R. (2000). *Highway Capacity Manual.* Transport Research Board.

[Coyle, 2007] Coyle, M. (2007). Effects of payload on the fuel consumption of trucks. *UK Department of Transport Report.* https://imise.co.uk/wp-content/uploads/2017/09/RR5-Effects-of-Payload-on-the-Fuel-Consumption-of-Trucks.pdf.

[Errampalli et al., 2015] Errampalli, M., Senathipathi, V., and Thamban, D. (2015). Effect of congestion on fuel cost and travel time cost on multilane highways in india. *International Journal for Traffic and Transport Engineering, 2015, 5(4): 458 - 472.* http://dx.doi.org/10.7708/ijtte.2015.5(4).10.

[Fagerholt, 2001] Fagerholt, K. (2001). Ship scheduling with soft time windows: An optimisation based approach. *European Journal of Operational Research, 131(3), 559-571.* https://doi.org/10.1016/S0377-2217(00)00098-9.

[Fagerholt et al., 2010] Fagerholt, K., Norstad, I., and Laporte, G. (2010). Reducing fuel emissions by optimizing speed on shipping routes. *Journal of the Operational Research Society, 61(3), 523–529*. https://doi.org/10.1057/jors.2009.77.

[Fan et al., 2019] Fan, H., Yu, J., and Liu, X. (2019). Tramp ship routing and scheduling with speed optimization considering carbon emissions. *Sustainability, 11(22),6367*. https://doi.org/10.3390/su11226367.

[Gentili and Mirchandani, 2012] Gentili, M. and Mirchandani, P. (2012). Locating sensors on traffic networks: Models, challenges and research opportunities. *Transportation Research Part C: Emerging Technologies, Volume 24, Pages 227-255*. https://doi.org/10.1016/j.trc.2012.01.004.

[Greenshields et al., 1935] Greenshields, B., Bibbins, J., Channing, W., and Miller, H. (1935). A study of traffic capacity. *Highway Research Board Proceedings, Volume: 14, Part 1, page 448-477*. https://trid.trb.org/view/120649.

[Hamad and Kikuchi, 2002] Hamad, K. and Kikuchi, S. (2002). Developing a measure of traffic congestion: Fuzzy inference approach. *Transportation Research Record: Journal of the Transportation Research Board, Volume 1802, Issue 1, pp. 77-85*. https://doi.org/10.3141/1802-10.

[Hickman et al., 1999] Hickman, J., Hassel, D., Joumard, R., Samaras, Z., and Sorenson, S. (1999). Methodology for calculating transport emissions and energy consumption. *Transportation Research Laboratory Report, No: SE/491/98*. https://www.academia.edu/12203823/Methodology_for_calculating_transport_emissions_and_energy_consumption.

[Hu et al., 2009] Hu, S., Peeta, S., and Chu, C. (2009). Identification of vehicle sensor locations for link-based network traffic applications. *Transportation Research Part B: Methodological, Volume 43, Issues 8–9, Pages 873-894*. https://doi.org/10.1016/j.trb.2009.02.008.

[Jereb et al., 2017] Jereb, B., Kumperščak, S., and Bratina, T. (2017). The impact of traffic flow on fuel consumption increase in the urban environment. *FME Transactions 46(2):278-284*. https://www.researchgate.net/publication/323249700_The_impact_of_traffic_flow_on_fuel_consumption_increase_in_the_urban_environment.

[Kalinic and Krisp, 2019] Kalinic, M. and Krisp, J. (2019). Fuzzy inference approach in traffic congestion detection. *Annals of GIS, 25(4),329-336*. https://doi.org/10.1080/19475683.2019.1675760.

[Kramer et al., 2015a] Kramer, R., Maculan, N., Subramanian, A., and Vidal, T. (2015a). A speed and departure time optimization algorithm for the pollution-routing problem. *European Journal of Operational Research, 247(3), 782-787*. https://doi.org/10.1016/j.ejor.2015.06.037.

[Kramer et al., 2015b] Kramer, R., Subramanian, A., Vidal, T., and Cabral, L. d. A. F. (2015b). A matheuristic approach for the pollution-routing problem. *European Journal of Operational Research, 243(2), 523-539*. https://doi.org/10.1016/j.ejor.2014.12.009.

[Kukadapwar and Parbat, 2015] Kukadapwar, S. and Parbat, D. (2015). Modeling of traffic congestion on urban road network using fuzzy inference system. *American Journal of Engineering Research (AJER), Volume 4, Issue 12, pp. 143-148*. https://www.researchgate.net/publication/289469971_Modeling_of_traffic_congestion_on_urban_road_network_using_fuzzy_inference_system.

[Mamdani, 1974] Mamdani, E. (1974). Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers, Volume 121, Issue 12, p. 1585 – 1588*. https://doi.org/10.1049/piee.1974.0328.

[McKnight et al., 2003] McKnight, C. E., S., L. H., Ozbay, K., Kamga, C., and Paaswell, R. E. (2003). The impact of traffic congestion on bus travel time in northern new jersey. *Transportation Research Board*. https://rits.rutgers.edu/files/bustraveltime.pdf.

[Norstad et al., 2011] Norstad, I., Fagerholt, K., and Laporte, G. (2011). Tramp ship routing and scheduling with speed optimization. *Transportation Research Part C: Emerging Technologies, 19(5), 853–865*. https://doi.org/10.1016/j.trc.2010.05.001.

[Rao and Rao, 2012] Rao, A. and Rao, K. (2012). Measuring urban traffic congestion - a review. *International Journal for Traffic and Transport Engineering (IJTTE) 2(4):286-305*. http://dx.doi.org/10.7708/ijtte.2012.2(4).01.

[Sung and Nielsen, 2019] Sung, I. and Nielsen, P. (2019). Speed optimization algorithm with routing to minimize fuel consumption under time-dependent travel conditions. *PRODUCTION & MANUFACTURING RESEARCH, 8(1), 1–19*. https://doi.org/10.1080/21693277.2020.1732848.

[Tišljarić et al., 2021] Tišljarić, L., Ivanjko, E., Kavran, Z., and Carić, T. (2021). Fuzzy inference system for congestion index estimation based on speed probability distributions. *Transportation Research Procedia, Volume 55, pp. 1389-1397*. https://doi.org/10.1016/j.trpro.2021.07.124.

[Toan and Wong, 2021] Toan, T. and Wong, Y. (2021). Fuzzy logic-based methodology for quantification of traffic congestion. *Physica A: Statistical Mechanics and its Applications, Volume 570, 15 May 2021, 125784*. https://doi.org/10.1016/j.physa.2021.125784.

[Treiber et al., 2007] Treiber, M., Kesting, A., and Thiemann, T. (2007). How much does traffic congestion increase fuel consumption and emissions? applying a fuel consumption model to the ngsim trajectory data. *Annual Meeting of* *the Transportation Research Board 2008*. https://www.akesting.de/download/How_Much_does_Traffic_Congestion_Increase_Fuel_Con.pdf.

[Wen et al., 2016] Wen, M., Ropke, S., Petersen, H., Larsen, R., and Madsen, O. (2016). Full-shipload tramp ship routing and scheduling with variable speeds. *Computers & Operations Research, 70, 1-8*. https://doi.org/10.1016/j.cor.2015.10.002.

[Wu, 2002] Wu, N. (2002). A new approach for modeling of fundamental diagrams. *Transportation Research Part A: Policy and Practice, Volume 36, Issue 10, Pages 867-884*. https://doi.org/10.1016/S0965-8564(01)00043-X.

[Xiao and Konak, 2016] Xiao, Y. and Konak, A. (2016). The heterogeneous green vehicle routing and scheduling problem with time-varying traffic congestion. *Transportation Research Part E: Logistics and Transportation Review, 88, 146-166*. https://doi.org/10.1016/j.tre.2016.01.011.