Fuzzy Logic Modelling of Route Choice in a Transportation Network

Sercan Malkoç
May, 2023

SUPERVISORS:

Dr. ir. Oskar Eikenbroek
Prof. Dr. ir. Eric van Berkum

ABSTRACT

Transport systems are one of the driving forces of economic and social development of societies. With the growing world economy and the increase in efficiency in manufacturing techniques over the years, automobiles have become affordable, and their number has increased rapidly. The increasing number of automobiles causes economic losses and social disturbances due to air pollution and noise pollution as well as traffic congestion. While these problems that arise in cities that cannot plan their infrastructure well are worsening day by day, the cost of the investment required for the solution of the problems that arise has been increasing. While developed countries aim to minimise the effects of increasing traffic-related problems by allocating more financial and human resources to plan urban transport infrastructures by taking into account the changing circumstances, developing countries face important problems such as economic and well-trained human resources in both planning and implementation of plans.

Understanding the route choices of road users in urban transport network planning is very important in terms of optimum use of the existing network capacity, understanding the deficiencies in the transport network, and deciding in which direction the infrastructure investments should be made. One of the most frequently used methods to understand the route choices of a road user is based on the comparison of the utilities of route alternatives to the road user. Quantifying the utility of a route to the road user is not easy since route choice is influenced by a wide range of factors, such as traffic safety, travel time, environmental factors, habits, etc. Logit and probit models are frequently used to estimate the utility of route alternatives. Although logit and probit models have various advantages, they have significant disadvantages in the application phase. Depending on the increase in the number of parameters considered in the logit model, the number of data should also increase in order to establish a reliable model. The probit model, on the other hand, requires mathematically complex operations, and its application requires experience and qualification. For this purpose, it is aimed to develop a model that can be used in small and medium-sized transport networks of developing countries by eliminating the disadvantages of logit and probit models, has low data dependency, is mathematically simple, linguistic expressions can be easily digitised, gives results quickly, and is easy to modify. The proposed model is a fuzzy logic model used to solve problems involving ambiguity. The fuzzy model is used in cases where a feature does not take a certain value and has applications in various engineering disciplines.

ACKNOWLEDGEMENTS

“Any place where you cannot go is not yours.” (Halil Rifat Pasha)

When I was an undergraduate student, our professor introduced transport engineering with this quote. For an undergraduate student, there probably would not have been a shorter phrase that could explain how crucial transport and logistics are. Life offered me the opportunity to follow the profound meaning of the quote thanks to the scholarship opportunity provided by the Ministry of National Education of the Republic of Turkey in 2018. I believe that the Logistics and Transportation track I completed at the University of Twente as part of this opportunity is an important milestone that will shape the rest of my life.

First of all, I would like to express my gratitude to the Ministry of National Education and the Ministry of Transport and Infrastructure of the Republic of Turkey for providing me with this scholarship and making a challenging but enjoyable journey of 2.5 years at the University of Twente. I would like to thank the Education Attaché of the Embassy in The Hague for ensuring that the scholarship programme has been carried out without interruption and for always making me feel their moral support throughout the scholarship program. Without these three actors and their support, neither this journey nor this study, which is the product of this journey, would have been possible.

I would like to thank Oskar Eikenbroek, who is my daily supervisor and who patiently listened to me, guided me, enriched the study with his feedback and weekly discussions, and taught me a lot. I would also like to thank Eric van Berkum for his interest and support throughout the period of finding my thesis advisor and for his feedback and suggestions after I started the thesis. I would like to thank all my professors with whom I crossed paths throughout the master's programme.

Çağlar Tabak, my advisor for the Turkey leg of the scholarship programme, has been one of my biggest motivators on this long journey by helping me overcome all the difficulties I faced during the master's programme by making me feel his trust and support in each conversation we had. I would like to specially thank him for always being contactable despite his busy work schedule and for the trust and support he has given me since the first day we met.

Lastly, I would like to thank my father Mustafa, my mother Nurgül, and my brother Ersel Malkoç, who have been with me throughout the whole exhausting journey and encouraged me to embark on this journey by always supporting me. In addition, endless thanks to all my friends, especially Şevket Feralan and Muhammet Ali Şeki, who supported me when I lost my faith in this long journey and helped me to finish it.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Failure to plan the transport network of a country, city, or region effectively has cost, time, and environmental impacts. A poorly designed transport network leads to an increase in the frequency of traffic jams and traffic accidents. Due to congestion and accidents as a result of poor design, road users are likely to spend more time, consume more fuel, and travel on a less safe road (Wang et al, 2013; Shi et al., 2016). All these factors have a direct impact on the quality of life of the people of the region, economically, socially, and environmentally. Therefore, proper planning of transport systems is of economic, environmental, and social importance.

There has been a significant increase in the number of automobiles in last decades as they have become more affordable thanks to technological and economic developments. The increase has also resulted in an excess demand for trips, which in turn creates new problems that need to be solved. As travel demand increases, it becomes more important for each driver to optimise their own cost and/or travel time. However, when users' self-interest optimisation reaches a saturation point, congestion occurs on a transportation network. Congestion on a traffic network has a direct impact on the travel time and cost of travellers. For this reason, various studies have been carried out on transport networks in order to alleviate traffic problems and ensure cost minimisation of users.

The shaping of a trip on a transport network is based on the decisions made by the users of the network. Each decision made by the road user is part of the transport planning process. In the transport planning process, the classical four-stage transport model is commonly used. This model consists of trip generation, which estimates how many trips are generated in the traffic network, trip distribution, which estimates where the trip starts and ends, modal split, which analyses how many trips are generated by which mode of transport, and traffic assignment, which assigns passengers to their routes according to their destinations. In this thesis, the concept of route choice, which is involved in traffic assignment, will be analysed.

Route choice in urban transport networks is a complex problem that cannot be handled in by considering the effect of one or two factors (Luce, 2005). Various factors such as traffic safety, congestion, distance, road works, travelling time, distance, environmental impacts influence the decision maker in route choice (Arslan & Khisty, 2005; Hawas, 2004) Typically, drivers aim to follow the shortest route to the destination they want to reach in the fastest way possible and without encountering any safety problems. It causes vehicle drivers to need an analysis process when choosing route alternatives that can meet their expectations at the highest level.

In a road network, the most basic approach to the route choice problem is generalised cost minimisation. Cost minimisation involves several different parameters, such as travel time, probability of road congestion, type of road, safeness of the road network and environmental impacts (Sheffi, 1985). It is exceedingly difficult to write a generalised cost function that takes all of these parameters into account and to construct a reliable mathematical model based on this function (Luce, 2005). In order to write a generalised cost function, it is necessary to write a cost value for the parameters affecting the route choice. Route length, travelling time, or the probability of road congestion can be expressed as a scalar value and can be converted into a cost. For this purpose, various approaches have been developed. These approaches essentially form utility functions that express the maximum benefit and cost provided by the routes. Based on these functions, logit or probit behaviour modelling methods are used, and the probability of choosing the routes is determined (Daganzo, 1979). However, the safeness of the road network or environmental impacts are not easy to express using a scalar value and difficult to convert to cost. Human reasoning is based on vague and subjective values rather than crisp numbers (Arslan & Khisty, 2005). Not everything in the human reasoning has precise values. In a crisp set, an expression is either 0 or 1, either yes or no, or more generally, either an element of the set or not. In human thinking methodology, an expression can belong to more than one set at the same time. For example, a route may be categorised as very safe by one road user due to appropriate lighting, pavement conditions, and traffic signs, while the same route may be considered as average safe by another road user for the same reasons. Therefore, making distinctions using sharp boundaries in human thought logic may not always give accurate results. In order to use the difference in the way of reasoning in the data evaluation procedure, fuzzy sets, which were introduced by Lotfi Zadeh (1965), are used as an alternative method. The use of fuzzy sets can enable the inclusion and analysis of parameters that are not considered, neglected, or cannot be separated with precise boundaries in the decision model. By utilising this approach, new fuzzy logic-based route choice models have been developed as an alternative to the classical model approach.

The route choice problem is a problem that needs to be addressed carefully before and after infrastructure investments. In order to make more efficient use of the existing infrastructure or to achieve the target of the investments to be constructed, it is necessary to determine the parameters affecting driver decisions and to analyse and understand them well. For the analysis of the route choice problem, it is assumed that travellers try to maximise their utility by choosing the most favourable alternative for them, which is generally defined as the most cost-effective route. This assumption is based on the theory of random utility maximisation Logit and probit models are the most common models for the explanation of the theory. Various advantages and disadvantages of these models are discussed under the chapter Route Choice Models. The main criticism of logit and probit models is that drivers' preferences cannot be captured accurately (Gärling et al., 1994). In addition, there is 60-80% similarity between the route choice prediction made with the overhead representation obtained by taking the weighted sums of the time and distance parameters, which are the two most important parameters affecting the route choice, and the actual observations (Outram & Thompson, 1977). Furthermore, gathering the necessary data for modelling and building a model on the basis of these data requires meticulous work in both the data gathering and modelling process. In this context, the inability to easily convert the effects of safety and environmental parameters into the cost function in the writing of the utility function, the difficulties in collecting qualified data and the need for qualified personnel in the evaluation of the collected data are the main deficiencies in the planning phase of small and medium-sized networks, especially in developing countries (Memon et al., 2009; Moubayed et al., 2009). In this context, due to its structure that allows easy quantification of linguistic expressions and its practicality of application, its suitability in network-wide route choice problems will be investigated.

1.1. Definition of the Problem

Route choice depends on various factors, such as travel time, traffic safety, environmental factors, traffic congestion, habits, and distance. It is very difficult to establish a route choice model by considering all of these parameters at once. The two most common models used for route choice models are probit and logit models. Both of these models are based on the estimation of the utility of the preferred alternative for the driver. The utility of the route alternative for the driver can be obtained by converting parameters such as travel time or distance, which can be expressed in scalar values, into travel cost and includes an error and perception term. However, it can be difficult to construct a utility function using parameters that are difficult to express using scalar values, such as traffic safety and environmental factors, and to observe the effects of these parameters on route choice. In addition, both logit and probit models have several disadvantages.

Probit model estimations involve multiple integral computations. As the number of alternatives increases, probit model estimations become more complex. The increasing number of alternatives both makes it difficult to estimate the distribution of the probability function and increases the labour required to solve the model. For this reason, the probit model is considered to be a suitable method for a set of alternatives with a maximum of three or four alternatives (Maddala, 1986). The logit model, another widely used model, also has several disadvantages. It requires a large data set to build a good prediction model using the logit model (Koppelman & Bhat, 2006). The number of data to be used increases as the number of variables in the model increases. The increase in the number of variables in the model negatively affects the predictive power of the logit model.

Both the complexity and limited predictive power of the probit model and the increasing data dependency according to the number of variables used in the logit model limit the use of both models in the analysis of small and medium-sized transport networks. In developing countries such as Turkey, traffic planning based on route choice by road users in small and medium-sized urban centres is not carried out generally. (Gençoğlu & Cebeci, 1999; Akintola et al., 2010). While one of the main reasons for this is the cost of planning projects, another important reason is that the job descriptions of engineers working in local governments cover various fields of civil engineering rather than having a job description only in a specific field such as traffic engineering (Memon et al., 2009; Moubayed et al., 2009). Therefore, instead of developing projects and specialising in a specific field, knowledge is acquired in various fields, and specialisation in one field is partial. In addition, another problem is that the data collected in small and medium-sized networks is quite limited.

1.2. Purpose of the Study

With the increase in the number of variables in the probit model, the model becomes more complex and requires expertise for the solution. In addition, the predictive power of the model decreases as the number of variables considered increases. Although the logit model is more widely used than the probit model due to its simpler structure and predictive power, the increasing number of independent variables also increases the data size to be used. It is aimed to develop an alternative model that can be used in route choice modelling by developing a simpler model rather than a model that requires qualification by eliminating the complexity of the probit model and by developing a prediction model with limited data by eliminating the data dependency of the logit model. For this purpose, the suitability of fuzzy logic, which can imitate human logic and is based on the rule-based expression of the relationships between the parameters considered, is easy to apply and has a wide range of applications, for the solution of the route choice problem will be investigated. The fuzzy logic model is intended to find the utility of a route to the road user without performing extensive calculations to construct a function as in the multinomial logit model, or minimising ambiguities in the categorisation of parameters whose effect on the cost function is expressed categorically, such as Traffic Safety and Environmental Factors. The study aims to develop an alternative model for engineers who cannot use one of the route choice models due to lack of financial resources, qualified personnel, or limited data, which will give them a basic idea for the route choice problem in their work areas and can be easily applied.

1.3. Importance of the Study

Fuzzy logic has been used in many fields for many years, both for its ability to process verbal information and for the expression of phenomena that are not numerical or cannot be easily expressed mathematically. One of the most prominent features of fuzzy logic is that it allows the building of models that can easily emulate human logic. Models that predict how human logic can make a choice according to the parameters defined by the rule sets to be specified can be built. Route choice is also based on road users making a choice as a result of considering various factors. Fuzzy logic will be used to investigate how travel time, environmental factors and traffic safety parameters affect route preferences. In this context, a fuzzy logic model based on these parameters will be built on the cycle route network in Enschede. In order to validate the model, open-source route data collected from volunteer cyclists by means of GPS and smartphone app within the framework of the "fietstelweek" event organised by Breda University Applied Sciences in 2016 will be used.

Although the first study based on fuzzy logic in route choice was carried out by Teodorovic and Kikuchi in 1990, studies on fuzzy logic and route choice are rarely encountered in the literature. Most of the studies on fuzzy logic in the literature concentrate on traffic assignment, which is the last stage of the four-stage traffic model. Murat and Uludağ (2008) and Dhulipala et al. (2020) developed fuzzy logic models that explain the route choices of road users through rule sets. In both studies, only one O-D pair was considered and alternative routes in the chosen O-D pair were evaluated through rule sets. This study aims to generalise and apply the fuzzy model on a network using rule sets, to find the probability of choosing any route among the alternatives with the obtained utility value, and to validate the established model using real world data.

1.4. Research Questions

Each link of the cycleway network in Enschede will be evaluated in terms of traffic safety, environmental factors, travel time and length. Then, the routes followed by the "fietstelweek" event participants will be compared with the shortest route between the origin and destination of the participants in terms of travelling time. Thus, these three different parameters will be used and inferred by building a fuzzy logic model, and the attractiveness for cyclists on the basis of route and link, in other words, the benefit will be obtained. In order to validate the model, the results found as a result of the fuzzy logic model will be compared with the route data of cyclists who participated “fietstelweek” event. The fuzzy logic will be tuned until the threshold, at least 80% of the routes followed have a better or close utility value than route alternatives, is satisfied. The questions to be answered within the scope of this thesis are as follows. Main question to be answered is,

“*How can a fuzzy logic model be built and validated to model route choice and explain route choices of cyclists using GPS data in a cycle network based on travel time, traffic safety, and environmental factors?*”

In order to address the main question, answers to the following sub-questions will be sought first.

The first sub-question is which sub-factors including traffic safety and environmental factors will be evaluated within the scope of this study. In this context, a literature review will be conducted on these two main factors. Following the literature review, it will be analysed which of the sub-factors for which data are available for Enschede. Thus, the sub-factors with data accessibility and emphasised in the literature will be included in the study. Considering all these, the first sub-question to be answered is,

"Which sub-factors related to traffic safety and environmental factors affect the route choice decisions of cyclists and how?"

By identifying the sub-factors affecting the main factors considered, traffic safety and environmental factors can be modelled with an "if-then" fuzzy logic rule structure approach to obtain the characteristics of all the links that make up the cycle network in the city of Enschede. The second question sought to be answered in this context is,

"How are the cycle network links in Enschede characterised in terms of traffic safety and environmental factors?"

After determining the features of the links according to the two main factors, they will be evaluated in terms of travel time, which is the main factor affecting route choice in many studies. Travel time will be evaluated in the context of the sub-factors obtained from the literature and the characteristics of each link in the cycle route network in terms of travel time will be obtained. In this context, the third question is,

"How are the links that make up the cycle network in Enschede characterised in terms of travel time?"

After identifying the characteristics of the cycle network links according to the three main factors, an analysis model can be built with "if-then" rule sets using fuzzy logic. One of the weaknesses of fuzzy logic is that it does not have a generalised systematic approach, and the rule sets and membership functions are based on experience and empirical. Therefore, it will be possible to understand the accuracy of both rule sets and membership functions by validating the model. This requires formulating the fourth question with two sub-questions.

"With what accuracy do the results obtained explain the GPS data obtained within the scope of 'fietstelweek'?"

"If the fuzzy logic model does not explain the results with sufficient accuracy, what is the main reason? Are the rule sets or membership functions in the fuzzy logic model faulty, or are the parameters/sub-parameters considered insufficient to explain the results?"

After the validation of the model, several interpretations can be made about the model constructed using fuzzy logic. The strengths and conveniences of the model as well as its failures and whether it is generalisable or not can be discussed. For this purpose, the fifth question can be formulated with two sub-questions.

"Can a fuzzy logic model be easily generalized? If no, what are the impediments to generalizability, and how can these impediments be overcome?"

"What are the main difficulties that may be encountered in the fuzzy logic model if more factors affecting the route choice are added? How might the addition of more factors affect the setup and accuracy of the fuzzy logic model?"

1.5. Assumptions and Delimitations of the Study

It is assumed that the participants of Fietstelweek - 2016, whose GPS data will be evaluated within the scope of the thesis, chose the routes they followed in order to maximise their benefits by acting in a rational human behaviour. The average speed of the participants is assumed to be 15.68 km/h. This is the average speed of all event participants in Enschede. Travel time is assumed to be influenced by motor vehicle and pedestrian traffic. Since there is very limited data on motor vehicle traffic, pedestrian traffic, and traffic light waiting time, several assumptions are made about their effects on travel time. Accordingly, the average time lost by a cyclist at an intersection with a traffic light is assumed to be 48 seconds in the city between 07:00-09:00 and 16:00-18:00 hours, when traffic is heavy and the waiting time changes dynamically depending on the traffic density. This time lost includes the time it takes to slow down for the traffic light, to wait until the green light is switched on, and to reach the former speed again. The time lost due to traffic lights is assumed to be 23 seconds outside the rush hour. In addition, on main roads that are not physically separated from motor vehicle traffic, it is assumed that cyclists may not be able to overtake the cyclist in front for safety reasons and reduce their speed in places during rush hours, and travel time is assumed to increase by 10%-15% on these links. Finally, it is assumed that cyclists reduce their speed by 20% for pedestrian safety due to the pedestrian traffic that occurs between 17:00-20:00 on weekdays and after 13:00 on weekends in the city centre where pedestrian traffic is heavy.

The most important limitation of the study is the date difference between the cycle network data to which the GPS data belongs and the Open Street Map (OSM) data from which the data such as land use, traffic lights, whether the bicycle lanes are separated from vehicle traffic are obtained. While the bicycle network that GPS data belongs to 2016, the OSM data belongs to 2023. In addition, it should be noted that OSM data is open source, and the data is created and checked by volunteers, so there may be various inaccuracies in the data or various discrepancies with the real data.

1.6. Organization of the Study

In the next section, the theoretical framework of route choice behaviour will be established and information about existing route choice models will be given. Then, the advantages and disadvantages of fuzzy logic will be discussed by giving information about fuzzy logic, which will be evaluated within the scope of the thesis and will be used to propose a new model. By giving information about the studies in which fuzzy logic is used to model route choice, the factors affecting the route choices of cyclists, which will be used in the study, will be discussed. In the third section, information about the methodology will be given. In the Methodology section, how the data used in the study are obtained, how the obtained data are prepared for the analysis environment, the theoretical framework of fuzzy sets and operations in fuzzy sets, and the construction of fuzzy logic rules will be presented to the reader. In addition, information on how to validate the established model will be given. In the fourth section, Results, the outputs of the model and the rule set derived from trial and error will be presented to the reader. In the following section, Discussion, the advantages and disadvantages of the fuzzy logic model analysed in this thesis will be discussed, and the research questions will be answered. In addition, suggestions for future studies will be given by emphasising how the fuzzy logic model discussed in the study should be improved. In the last chapter, Conclusion, the methodology followed in the study will be summarised and the findings of the study will be presented.

2. THEORETICAL FRAMEWORK

In this chapter, fundamental information about the basic concepts that will be discussed in the thesis will be given. Firstly, route choice behaviour will be discussed and detailed. Then, the existing theories used to model route choice will be discussed, and the strengths and weaknesses of the existing theories will be highlighted. This will be followed by general information about fuzzy logic, which will be discussed in the thesis, and which is assumed to overcome the weaknesses of the existing theories, and the proposed methods for solving the route choice problem with fuzzy logic. Finally, the factors affecting the route choices of cyclists will be discussed, and the purpose of the thesis will be referred to again in this context. Thus, it is aimed to explain the general concepts and theories related to the study and to provide a better understanding of the Methodology section following this chapter.

2.1. Route Choice Behaviour

Travelling is an important part of daily life and is shaped by the mode of transport and route choice. When travelling, road users evaluate the route alternatives according to themselves and typically choose the most beneficial one for them. Different studies on route choice show that route choice depends on various factors, such as travel cost, travel time, traffic safety, comfort, habits, and socioeconomic characteristics (Arslan & Khisty, 2005; Prato & Bekhor, 2007). Among these parameters, the most prominent factor is travel time (Bovy & Stern, 1990).

The most general approach to route choice, which is also the basis for traffic assignment, is that commuters evaluate the total costs of all routes on a route and anticipate and choose the route with the lowest total travel cost (Teodorovic & Kikuchi, 1990). It is argued that differences in the route choice of drivers travelling between two points are mainly due to two reasons (Bovy and Stern, 1990). The first reason influencing choice is human trait. Drivers have different intuitions when choosing a route (Bovy & Stern, 1990). The main factor influencing route choice behaviour is not the difference in the parameters that different researchers combine in their overall cost functions, but the difference in the perception of these parameters by drivers. The second reason for differentiating route choice behaviour is congestion effects due to unpredictable accidents or excessive demand (Bovy & Stern, 1990). Shorter routes are initially favoured more, so that congestion effects affect the travel time on these routes more than the initial conditions. In combination with the congestion effect, the cost of shorter routes, which are preferred because they are less costly, can increase to levels comparable to less favoured alternatives. Moreover, the preference of a route by more drivers statistically increases the probability of a traffic accident on the route (Ivan, 2004). It is not possible to know exactly whether a traffic accident will occur on the route. Other unknowns can be when the traffic accident occurs, how many vehicles will be involved in the traffic accident, the location of the traffic accident, and how the service capacity of the route will be affected depending on the number of vehicles involved in the accident. So, in the case of an accident, route costs may increase unpredictably and rapidly.

Parameters affecting route choice behaviour include travel time, traffic safety, traffic congestion, travel cost, road type, distance, traffic signing, road works, landscape, and driver habits (Arslan & Khisty, 2005; Hawas, 2004). It is very difficult to produce a generalised cost representation that includes all these parameters. It is obvious that it is not practical to represent all these parameters in a traffic assignment model. Therefore, various simplifying approaches for route choice models are inevitable. A common approach in route choice models is to consider the time and cost parameters. Especially when the urban traffic network is considered, time becomes the parameter to be considered (Bovy & Stern, 1990). Cost is often directly related to travel time (Bovy & Stern, 1990).

In the next section, the development of various models to explain route choice and the involvement of human behaviour in these models are discussed.

2.2. Route Choice Models

Differences in drivers' goals, expectations, and intuitions lead to differences in route choice behaviour, and different route alternatives are likely to be chosen, even if they are not the most attractive route. These behavioural differences can emerge as a stochastic element in the modelling. Stochastic approximation models have been developed to extend the restrictive assumptions in UE models and to develop a more realistic route choice behaviour model (Burrell, 1968; Dial, 1971). The developed stochastic models take into consideration the differences in drivers' intuition of costs, as well as the differences in the travel distance, travel time, and travel cost parameters that drivers aim to minimise.

Discrete choice models assume that when confronted with a choice situation, each individual's preference for alternatives is shaped by the attractiveness or utility criterion of the alternative (Cantillo & Ortúzar, 2005). The attractiveness of alternatives is represented by the concept of utility, which each individual tries to maximise. The value of utility is represented as a function of the characteristics of the alternatives and the characteristics of the decision maker. Since it is not possible to directly observe or measure the utilities, the parameters that are thought to benefit the individual are treated as random, and the route choice model shows which route will be chosen with what probability (Cantillo & Ortúzar, 2005).

Probit and logit models have been developed to model the influence of human behaviour on the route choice process more explicitly (Jones, 2021). Both models are based on the concept of utility in travel demand modelling. The attributes and criteria of the alternatives within the choice set confronting each decision maker are assumed to be defined values (Bunch et al., 1996). Drivers are thus regarded as rational decision makers who maximise their utility.

The utility that a road user *i* would obtain from an alternative route *a* is,

$$U_a^i = V_a^i + \varepsilon_a^i \tag{2.1}$$

In this equation, V is the deterministic component and ε is the random component used to account for uncertainty that represent the temper of each individual and measurement/observational errors made by modeller (Ortúzar & Willumsen, 2011).

Decision makers may not have complete information about the road network, and they cannot forecast the uncertainties or not accurately. As a result, predicted cost is affected by these uncertainties. Therefore, the uncertainties that may be encountered should be taken into account, and the utility function should be rewritten to incorporate these uncertainties. The sources of errors are based on the theory of random utility, which argues that some components of utilities cannot be observed by the modeller and should therefore be considered random (Cascetta, 2009). This variation in random utility models is directly dependent on the assumptions made for the random component ε and/or the deterministic component V of the equation (Cascetta, 2009). The probability of choosing an alternative is expressed as follows.

$$P_a^i = \Pr\left(U_a^i \geq U_b^i\right)$$

$$P_a^i = \Pr\left(V_a^i + \varepsilon_a^i \geq V_b^i + \varepsilon_b^i\right) \tag{2.2}$$

$$P_a^i = \Pr\left(V_a^i - V_b^i \geq \varepsilon_b^i - \varepsilon_a^i\right) \forall\, a \neq b$$

P_a^i is the probability that alternative a will be chosen under the current conditions, i is a road user, and it is stated that the maximum utility will be achieved if alternative a is chosen among all alternatives in the choice set. Commonly used random utility models are probit model, logit model, and generalised extreme value (GEV), discussed in Sections 2.2.1, 2.2.2, 2.2.3, and 2.2.4, respectively.

2.2.1. Multinomial Probit Model

The multinomial probit model fitted to the normal distribution is constructed in a similar way to the multinomial logit model, discussed in Section 2.2.2, fitted to the Gumbel distribution. However, the multinomial probit model uses the cumulative normal distribution function instead of the cumulative logit distribution function (Paleti, 2018). This is due to the assumption that the error terms are normally distributed. Unlike the multinomial logit model, this model does not assume that the error terms are uncorrelated (Brooks, 2008). In other words, a positive correlation between the error terms to reflect a similarity in the characteristics of two or more alternatives is not an impediment to the use of the multinomial probit model (Brooks, 2008).

The multinomial probit model is appropriate when there is a limited set of alternatives. The model is suitable for the choosing of at most three or four different alternatives (Maddala, 1986). Since the calculations of the probit model require the use of multiple integrals, the complexity of the calculations becomes more intricate as the number of alternatives increases.

The multinomial probit model, despite its theoretical attractiveness, considerable flexibility and advantages, is not a frequently used method. This is mainly due to computational difficulties. The first computational difficulty is that maximum likelihood estimation of complex nonlinear models is difficult for many practitioners to model (Bunch & Kitamura, 1989). Thus, the multinomial probit model is more complex than many other discrete choice models. Another difficulty is that more useful specifications require the estimation of covariance parameters, and the properties of the probability function are almost unknown in this context (Kropko, 2007). Finally, choice probabilities may require integration of the multivariate normal probability density. The workload required for standard integration approaches increases exponentially with the number of alternatives in the chosen set (Bunch & Kitamura, 1989).

2.2.2. Multinomial Logit Model

In the multinomial logit model, the probability distribution is S-shaped. If the representative utility of an alternative is very low or very high compared to other alternatives, a small increase in utility occurs. The point at which this increase in the representative utility of an alternative has the greatest effect on the probability of being chosen is when the representative utility is equal to the combined utility of the other alternatives (Koppelman & Bhat, 2006). Accordingly, a small increase in the utilisation of an alternative can determine the equilibrium and lead to a large increase in the probability of choosing the alternative (Koppelman & Bhat, 2006).

In the model, there is no order in the categories of the dependent variable, and the error terms are independent and constant variance. The probability of choosing an alternative i from a set of J alternatives is shown in Equation 2.3. In the equation, $P(i)$ is the probability that the decision maker chooses alternative i, and V_i is the systematic utility component of alternative i. $P(i)$ is always greater than zero for all i alternatives on the network.

$$P(i) = \frac{e^{V_i}}{\sum_{i=1}^{J} e^{V_i}} \tag{2.3}$$

One of the most controversial features of the multinomial logit model is the independence of irrelevant alternatives (IIA) (Cheng & Long, 2007). IIA means that the probability rate of any decision maker's choice between two alternatives is independent of the existence or characteristics of other alternatives. In other words, other alternatives are unrelated to the decision to choose between two alternatives within the set. The property has some important implications for the use, formulation, and estimation of the model. IIA allows an alternative to be added or removed from the choice group without affecting the structure and parameters of the model (Cheng & Long, 2007). However, the flexibility in applying the model to conditions with different choice alternatives leads to various advantages. The first of these advantages is that the model can be applied to situations where different members of the sample have different alternatives. Secondly, IIA facilitates the estimation of parameters in the multinomial logit model. Finally, when a new alternative is added to the choice model, it is advantageous in estimating the probability of choosing of this alternative (Koppelman & Bhat, 2006).

2.2.3. Nested Logit

The IIA assumption of the multinomial logit model significantly restricts the model as it requires equal competition between all pairs of alternatives, which is an inappropriate assumption under conditions where there are many alternatives (Björnersedt & Verboven, 2014). Thus, the ratio of the choice probabilities of any pair of alternatives is constrained to be independent of the existence and properties of other alternatives in the choice set. This constraint implies that the introduction of a new mode or improvements to any existing mode reduces the choice probabilities of the existing modes in proportion to their probabilities before the change. For example, suppose that in the case of urban transport mode choice there are alternatives of private car, bus, and light rail. Due to shared attributes not included in the measured portion, the bus and light rail alternatives are likely to be similar to each other relative to the other binary choice alternatives. However, IIA states that these alternatives are not similar. Therefore, there are cases where the IIA assumption does not appropriately reflect the behavioural relationships between alternative groups (Koppelman & Bhat, 2006).

Although it is easy to work with the multinomial logit model and to estimate with the model, various estimation errors occur when the IIA, which is a significant assumption for the reliability of the analysis results, is met (Carrasco & Ortúzar, 2002). Since the probit model, which was the only alternative when this assumption was not met until the 1980s, could not respond with sufficient flexibility to the needs when the choice set was more than three, McFadden (1981) proposed GEV models. Practical applications of GEV models have created a subclass called nested logit. The nested logit model, first derived by Ben-Akiva & Lerman (1985), is an extension of the multinomial logit model designed to capture the correlation between alternatives. The nested logit model is a partitioning of a choice set (C) into nests (C_k, C_j...). For each pair $C_k \cap C_j = 1$. For each alternative, the utility function consists of a part associated with the alternative and a part associated with groups of alternatives.

The nested logit model is represented by a tree structure in which similar alternatives are grouped into sets with no commonalities between them. This model is considered as a more flexible and statistically improved version of the multinomial logit model. With this model, similarities between overlapping alternatives can be captured, and the model offers a more flexible error structure (Carrasco & Ortúzar, 2002). In the nested logit model, an alternative can belong to only one set. Since the choice set or alternatives are divided into many sets, the model becomes very complex in a real road network where a link may belong to more than one set (Lai & Bierlaire, 2015). The model does not take into account the correlation between sets. It performs worse than the probit model in capturing partial overlap on urban roads (Gommers & Bovy, 1986).

2.2.4. C-Logit

C-logit is a model developed by adding a commonality factor to the utility function of the multinomial logit model. The added commonality factor represents the links shared between overlapping routes (Zhang & Du, 2020). Since the commonality factor is derived from the utility of overlapping routes, the utility of overlapping routes decreases while the utility of independent routes increases (Zhang & Du, 2020).

The probability of choosing any route *n* in the presence of more than two route alternatives is generally expressed by Equation 2.4.

$$P_n = \frac{e^{V_k^n - cf_k}}{\sum_{h \in I_{rs}} e^{V_h^n - cf_h}} \tag{2.4}$$

In this equation, cf_k is the commonality factor of route k and is directly proportional to the degree of commonality between route n and the other routes in the O-D pair. Cascetta et al. (1996) expresses the cf_k term as in Equation 2.5.

$$cf_k = \beta_0 ln \sum_{h \in I_{rs}} \left(\frac{L_{hk}}{L_h^{0.5} L_k^{0.5}} \right)^{\gamma} \tag{2.5}$$

Here, L_{hk} is the length of links common to paths h and k, L_h and L_k are the total lengths of routes h and k, respectively, and γ is a positive parameter.

2.3. Limitations of the Current Models

Route choice depends on different factors, such as travelling time, traffic safety, environmental factors, traffic congestion, habits, and distance. Considering all of these factors together makes a route choice model complex and intricate to solve the problem (Arslan & Khisty, 2005). The route choice models mentioned above are the most common methods used for route choice. New methods have been derived to overcome the shortcomings of the logit model. Some of the disadvantages of the logit model have been overcome by these new logit models. Although the new models explain route choices more accurately, one of the biggest disadvantages of the logit model derivatives is that they are quite complex compared to the standard logit model and require more data for the estimation of the model and more assumptions for the distribution of these data (Cascetta et al., 1996; Koppelman & Wen, 2000). Since the probit model is based on multiple integral calculations, the workload required for the probit model accrues as the number of alternatives increases. The probit model, which becomes more complex with the increase in the number of alternatives and the estimation of probability functions becomes difficult, is a suitable method in cases where a choice between three or four alternatives is required (Maddala, 1986). In short, the predictive power of the logit model and its derivatives decreases as the number of variables considered increases and the number of data required for the solution increases and the solution becomes more complex. In the probit model, as the number of alternatives increases, the required processing load increases, and it becomes difficult to estimate the distribution of probability functions.

2.4. Fuzzy Logic and Scope

Route choice on a network is an intricate problem due to the uncertainties involved. The fact that road users choose the routes they are used to or that they do not have sufficient information about various route options may cause various problems in the models established while addressing the route choice problem. For example, the experience of the road user, the fact that the same route has different travel times at different hours of the day, or on different days of the week cause various uncertainties that need to be addressed in the problem. The random component of the problem can be modelled with the aim of maximising the utility with various approaches mentioned above. However, one of major disadvantages of these methods is the necessity to study with precise values. The uncertainties, for example travel time affected by many factors, in the structure of the problem cannot fully respond to the uncertainty in the structure of the problem if precise values are used (Henn, 2000). In addition, in order to construct a realistic route choice model by simplifying the complexity of the route choice problem, the imprecision, vagueness, and ambiguity properties of the parameters in the route choice model must be understood correctly (Lotan & Koutsopoulos, 1993). Unlike classical mathematical models, fuzzy logic makes it possible to model these properties.

Classical logic deals with propositions with crisp values that occur only under certain conditions. However, in real life, the number of situations that we can easily distinguish as true or false is quite limited and does not occur based on absolute distinction. In addition, the requirement that the data used in classical logic must be precise causes various difficulties in solving the problems addressed in real life (Zadeh, 1965). Fuzzy logic, developed to overcome these difficulties, can be defined as a flexible application of logic rules (Zadeh, 1965). Classical logic is bivalent Aristotelian logic and contains only values, such as "true or false", "1 or 0", "yes or no", "exists or does not exist". Fuzzy logic, on the other hand, generalises classical logic and includes propositions and statements that can take any value between two values. Accordingly, "fuzziness" means multi-variable mathematically. Therefore, it is more effective in solving problems where uncertainty or truth values are not precisely defined, and the application area of fuzzy logic is wider (Trillas & Eciolaza, 2015). The mentioned comparison of "bivalence" and "multivalence" is one of the principles of fuzziness. The fact that real-life problems are predominantly multi-valued makes the bivalent approach of classical logic inadequate, and it is known that the use of fuzzy logic is more appropriate (Trillas & Eciolaza, 2015). Fuzzy logic theory is often confused with probability theory. These two theories have different definitions. While probability theory deals with uncertainties about whether a well-defined event occurs or not, fuzzy logic theory deals with uncertainties in defining the phenomenon (Trillas & Eciolaza, 2015).

Fuzzy logic has several advantages due to its flexible structure. The first of these advantages is the effectiveness of fuzzy logic in processing verbal information. The human mind uses words. Since fuzzy logic has similar properties, inferences close to human logic can be obtained with this method (Albertos et al., 1998). Another advantage of fuzzy logic is that it allows modelling of experiences that are not numerical or cannot be expressed mathematically (Mendel, 1995). For example, a parent who monitors his/her children closely can predict how many pages will be read each day from the books he/she gives to his/her children to read. However, predicting the number of pages that will be read can be difficult in terms of formulation. In this respect, it is a good tool for transferring experiences to the problem solution model in situations where numerical data cannot/limitedly be reached. Fuzzy logic includes the mathematical concepts of set theory, and the logic behind of the method is quite simple. With its flexible structure, it can keep in harmony even with complex and non-linear functions in uncertain probabilistic cases (Albertos et al., 1998). Colloquial language can be emulated by using fuzzy logic. It is therefore a tool to help build a predictive model for situations where it is difficult to build a mathematical model or where there are not enough data/human resources to build a complex mathematical model (Safiotti, 1997). Finally, fuzzy logic also includes classical logic and the phenomena constructed with classical logic can also be expressed with fuzzy logic (Albertos et al., 1998). In addition to all these advantages of fuzzy logic, it also has disadvantages. The most important of these is that there is no systematic approach to be followed for the solution of the problem with fuzzy logic. The choice of membership functions is very dependent on experience, and the most appropriate membership values of functions are found empirically (Albertos et al., 1998).

2.5. Fuzzy Logic and Application to Route Choice Problems

Fuzzy logic is a method that acknowledges imprecision, vagueness, and ambiguity in accordance with the nature of human thought. Unlike classical logic methods, this method works with values expressed in scales instead of a limited and explicit language expressed in numbers. The values on the scales need not be strictly separated from each other. For example, a journey time of about 15 minutes may be characterised as short for some road users and as average or long for others under the same conditions. The route choice problem is characterised by various uncertain attributes, such as travel time, distance, comfort, safety, etc. that road users assign to route characteristics. Fuzzy logic allows the input variables to be expressed linguistically, whereas models based on precise choices do not have the capacity to incorporate the uncertainty and ambiguity that dominate route choice decisions. The method is based on a set of rules generated by a simple "if-then" query. The route choice problem can be expressed as "Route A is both shorter and safer than route B, so route A is more likely to be preferred by a road user than route B". Thus, all factors to be considered on the route can be expressed in a rule-based approach and routes can be realistically evaluated in terms of their qualitative characteristics.

Teodorovic and Kikuchi (1990) presented the first study using a fuzzy logic model to solve the route choice problem. They developed a rule set based on the travel time difference between two routes. The travel time difference between two routes was expressed as fuzzy with four different categories and a fuzzy logic probability rule set was developed that generates a value between 0 and 1 according to the result of the travel time difference output. As a result of the defuzzification of the developed rule sets, the choice probabilities of the two routes analysed were calculated. Lotan and Koutsopolous (1993) introduced a new method based on the fact that the fuzzy logic model developed by Teodorovic and Kikcuhi is limited to a binary choice and the proposed method cannot be generalised easily for multiple choices. In the model developed by the duo, it is accepted that the most important factor affecting the route choice is the travel time and a fuzzy logic model based on the travel time is built. The travel time advantage/disadvantage of the considered route over other routes is quantified with the defined rule set fuzzy logic model. The quantified values are summed and defuzzified and the attractiveness of each route according to travel time is obtained numerically. Vythoulkas and Koutsopoulos (2003) proposed a new approach by extending the fuzzy logic framework developed by Lotan and Koutsopoulos to include the weights of the defined rules that form the decision process. In the study, fuzzy sets and verbal variables are used to model how the decision maker perceives different attributes (time, cost, and transfers) of alternatives. The basic approach in the model is that the decision taker makes a decision based on a few simple inferences with these attributes. These inferences are expressed as rule sets using fuzzy logic and artificial neural networks are used to find and calibrate the weights of each rule. The basic hypothesis on which the model is based, that drivers will decide based on simple rules rather than trying to maximise their decisions by writing a complex utility function, is tested. For this purpose, a logit model was developed, and the results of the proposed model are compared with the logit model. The proposed model gives better results compared to the logit model.

Murat and Uludağ (2008) investigated the probability of choosing four different alternative routes between a chosen O-D pair in Denizli/Turkey by using fuzzy logic and logit model with a survey of 500 respondents using the parameters of travel time, traffic safety, environmental impacts, and road congestion. They developed a rule-based fuzzy logic model consisting of three categories for each parameter. With the defined fuzzy logic rule model, a utility function was obtained according to the characteristics of each route, and the obtained utility function was converted into a probability value using a logit model. The survey results are also quantified using a logistic regression model and the probability of choosing each route is found. The fuzzy logic and logistic regression models are compared, and it is stated that the route choices of the respondents are better explained by the fuzzy logic model. In a similar approach, Dubey et al. (2013), based on a survey of 150 respondents in Delhi, modelled the probabilities of choosing three different route alternatives between an O-D pair using three different methods by using nine different variables including distance, speed, time, and delay. The authors compare the results obtained by using multinomial logit (ML), fuzzy rule-based inference system (FIS), and adaptive neuro-fuzzy inference system (ANFIS) models. As a result of the rule sets defined in the FIS model, a route preference value is obtained. The obtained value is converted into a probability value by using the multinomial logit formula. The results for all three models are similar to each other. The models have predicted the data set with an accuracy in the range of 95%-100%. The authors emphasised that the FIS model is the most effective method in terms of ease of calculation and flexibility. A study similar to these two studies was conducted by Dhulipala et al. (2020) in Surat, India. Three different route alternatives between an O-D pair in Surat are evaluated in terms of travel time, traffic density, and environmental impact parameters. The choice probabilities of each route alternative are calculated using multinomial logit model (MNL) and fuzzy logic. Unlike the other two similar papers, the authors do not utilise the multinomial logit formula to calculate the probability of choosing a route and propose a hierarchical method. Accordingly, the defined rule sets are defuzzified as a number ranging from 0-100 for the route choice preference of three different parameters. According to the drivers' answers to the questionnaires, three different routes are compared and grouped in groups of two. For example, when comparing routes, A and B, if a road user's preference for route A is obtained as a value of 50 or more than the preference for route B, it is assumed that the driver chooses route A between route A and route B. In the next stage, the comparison is made for routes A and C in the same way to find out which route the driver ultimately prefers. According to this validation model developed by the authors using fuzzy logic, the survey data is explained with 92% accuracy, while the model built with MNL achieved to explain the survey data with 78% accuracy.

2.6. Route Choice Parameters of Cyclists

Unlike motor vehicle drivers, cyclists consider many incomparable objects when making a route choice decision. It is not realistic to consider generalised cost as the only objective cyclists aim to achieve (Ehrgott et al., 2012). In this respect, cyclists might differ from motor vehicle drivers. Cyclists have different factors affecting their route choice decisions. Therefore, unlike motor vehicle drivers, it is essential to understand what factors cyclists consider for route choice other than travel time or distance in order to make a good route choice prediction and to be able to model it mathematically.

Many researchers have conducted numerous studies centred on cyclists' route choices. Based on the literature, the findings can be categorised into three main categories: traffic safety, environmental factors, and travel time. Although the main purpose of transportation is to get from one origin to another destination as fast as possible, this purpose cannot be considered in isolation from safety. Especially for cyclists, the shortest route does not always mean the safest route. Cyclist safety is only partially recognised where cycle lanes alongside main roads are not separated by a physical barrier (Ehrgott et al., 2012). If the cyclist has an alternative that is much safer and does not have a significant difference in terms of travel time, the cyclist may choose this route alternative for safety reasons. Studies in the literature confirm the cyclists' view of safety. Hopkinson and Wardman (1996) conducted a survey of 7900 respondents in Bradford, West Yorkshire in 1994 showed that cyclists place more importance on safety than travel time. In addition to traffic safety, the effects of environmental factors on cyclists' route choice have been confirmed by various researchers. Various factors, such as the number of traffic lights and intersections on the route, variation in elevation, water bodies, land use diversity, public transport transfer on the route, the presence of a facility such as a gym, shopping centre, the presence of green space can be a reason for cyclists' route choice (Prato et al., 2018; Hood et al., 2013; Menghini et al., 2010). In addition, some of these factors have a direct and some have an indirect effect on travel time. In the following subsections, the studies and findings in the literature on the effects of travel time, traffic safety, and environmental factors on cyclists' route choice are presented.

2.6.1. Travel Time

Travel time is considered to be the most important element of the utility value used to find the probability of choosing a route, and most modelling based on route choice is based on this assumption. As emphasised before, although the parameters that are valid for route choice for cyclists are different from motor vehicles, it is undeniable that travel time is also one of the main factors in route choice for cyclists. It is not possible to consider travel time independently from traffic safety and environmental factors. Travel time is affected by both safety factors on the chosen route and environmental factors. For example, due to the presence of residential areas along the route or areas with a high pedestrian population, such as a city centre, may require cyclists to ride at a slower than normal speed for their own safety and safety of pedestrians. Or, the presence of traffic lights on a route may increase safety for cyclists at intersections, while waiting at traffic lights lengthens the travel time.

The presence of a cycle path affects traffic safety as well as travel time. On a cycle path separated from motorised traffic, cyclists can travel as fast as their physical condition allows without safety concerns (Clarry et al., 2019). Band (2022) emphasises that cyclists are willing to cycle on average 3.3 minutes more to travel on a route with a cycle lane. The same study also states that it is not a problem for cyclists to travel 4.7 and 5.2 minutes longer in the presence of a one-way or two-way cycle track, respectively. Sobhani et al. (2019) state that cyclists tend to use routes where they can avoid traffic as much as possible during peak traffic hours. In addition, many factors such as the number of intersections on the route, the number of red lights, left turns, the quality of the road surface and the road slope play an important role in both the travel time and the route preference of cyclists (Lawrence & Oxley, 2019; Lu et al., 2018; Ton et al., 2017; Beheshtitabar et al., 2014). Due to these factors, the shortest road alternative in terms of distance may not be preferred by cyclists.

2.6.2. Traffic Safety

A common finding of research on cyclists' route preferences is that routes that offer a safe infrastructure for cyclists are more attractive routes for cyclists. Cyclists are vulnerable road users and in the event of a vehicle-cyclist collision, the cyclist is more likely to be seriously injured or die than the vehicle driver (Klanjčić et al., 2022). Therefore, separating cyclists from vehicle traffic with a physical barrier, especially on roads with a speed limit of more than 50 km/h, is important to ensure road safety for cyclists. The preference and use of cycle lanes separated from motorised traffic by a physical barrier is increasing significantly (Trofimenko & Shashina, 2021; DiGioia et al., 2017; Chen et al., 2016). In cases where physical separation of the cycle lane is not possible due to infrastructural deficiencies, clearly marking and painting the cycle lane is an effective and inexpensive method to prevent vehicle drivers from violating the cycle lane while driving (Oršić et al., 2022). Similarly, the marking of intersections for cyclists is a simple and effective practice for the navigation and safety of both vehicle drivers and cyclists at an intersection (Schepers et al., 2011; Wall et al., 2016).

The characterisation of a cycle route as safe is not limited to its physical separation from vehicular traffic or the marking of the cycle route to make it easier for everyone to navigate. The superstructure of the cycle path is another factor to be considered for a safe and comfortable cycling experience and the attractiveness of the route. The presence of potholes on the surface of the cycle path or the deterioration of part of the road directly affects cyclist safety (Dondi et al., 2011). Potholes on the road may cause the cyclist to trip and fall while travelling and affect the road grip. In addition, the pavement material used on the cycle path should provide sufficient friction to ensure grip and safe riding on wet surfaces (Dondi et al., 2011). Finally, lighting on the cycle path is considered as a factor affecting cyclist safety for night journeys. Illumination of cycle paths increases cyclist safety by increasing the visibility of cyclists, enabling faster action in case of any factors affecting driving safety, and providing a mentally more comfortable driving experience (Osama & Sayed, 2017). The most comprehensive study on this subject was conducted by Wanvik (2009) in the Netherlands and concluded that road lighting in rural areas reduces the number of accidents by half, as well as the severity of the accidents that occur. In addition, various researchers have shown that lighting reduces the likelihood of cyclists being involved in an accident (Osama & Sayed, 2017; Chen & Shen, 2016; Reynolds et al., 2009).

2.6.3. Environmental Factors

Unless the cycle route is independent of motorised traffic, any factor affecting motorised traffic can also affect cyclists. One of the most important environmental factors for cyclists' route choice is the number of stops along the route (Lawrence & Oxley, 2019; Hood et al., 2013; Menghini et al., 2010). The red-light duration of traffic lights depends on the volume of vehicle traffic due to the widespread use of smart signalling systems in many cities. Therefore, for a route choice with heavy traffic, the waiting time at a red-light may be longer than expected. Signalling at intersections has a negative impact on travel time compared to traffic safety. Several studies show that the presence of signalling on the route affects cyclists' route preferences differently. While Ehrgott et al. (2012) and Lawson et al. (2013) argue that traffic signalisation increases cyclists' perception of safety and makes the route preferable, Lawrence

& Oxley (2019), Ton et al. (2017), and Hood et al. (2013) conclude that traffic signalisation affects travel time and negatively affects cyclists' route preferences. In a study by Band (2022), cyclists prefer to travel 5.3 minutes more on a route with two or fewer stops rather than stopping five or more times on a route due to signalling or intersections.

Another factor affecting the route preferences of cyclists is the road gradient. Gradient of a cycle path is a negative factor for cyclists as they require more effort than flat paths. Scarf & Grehan's (2005) study shows that cyclists prefer to travel 8 m horizontally rather than 1 m vertically. Cyclists do not tend to detour on routes with relatively low road gradient, while the tendency increases in terms of gradient. The study of Lu et al. (2018) emphasises that a road gradient of 2% or less makes the route an ideal route alternative for cyclists. The same study states that road gradients between 2% and 6% make the route partially detouring for cyclists, while road gradients above 6% make the route highly detouring for cyclists (Lu et al., 2018).

Cyclists' route preferences are influenced by road characteristics and safety as well as environmental factors, such as the presence of green space, blue space, recreational space, parkland on the route, which make the cycling experience more enjoyable and relaxing. Prato et al. (2018) emphasise that scenic areas contribute positively to cyclists' route preferences. Marquart et al. (2020) emphasise that the presence of blue space on the route has a positive effect on the cycling experience and that cyclists with no time constraints are more likely to choose such cycling routes by sacrificing travel time if they have a longer travel time. In addition to green and blue areas, the presence of facilities on the road also affects the route choice. The presence of facilities with different qualities, such as grocery market, shopping centre, bakery, library, gym/fitness facility, school, recreation facility on the route has a positive effect on the route choice decisions of cyclists (Kerr et al., 2016).

Finally, the diverse effects of land use and demographics on cyclists' route choice and route attractiveness can be addressed. Cycling in residential areas does not allow for a smooth cycling experience due to the relatively narrow roads, the lack of separated cycle lanes, and the presence of many uncontrolled intersections with crossing side streets. Therefore, cyclists avoid cycling in residential areas and densely populated settlements as much as possible (Koch & Dungundji, 2021; Prato et al., 2018). Studies investigating the impact of mixed land use, industrial land use, and commercial land use on cyclists' route preferences are limited in the literature. Winters et al. (2010) and Prato et al. (2018) emphasised that mixed land use has a positive effect on cyclists' route choice, and Prato et al. (2018) indicated that industrial land use contributes positively to cyclists' route choice. Koch & Dugundji (2021) state that commercial land use has a positive effect on route choice, while Winters et al. (2010) argue that commercial land use has a negative effect on route choice with the opposite finding.

2.6.4. Overview of Factors

In the sub-headings above, several factors affecting the route choice behaviour of cyclists are grouped and discussed as travel time, traffic safety, and environmental factors and supported by the studies in the literature. Accordingly, traffic safety plays an essential role in cyclists' route choice behaviour. A cycle lane physically separated from vehicular traffic increases the safety perception of cyclists. Cycle lanes marked with a different colour than the road pavement at intersections and non-physically separated cycle lanes have an importance for the navigation of both cyclists and motor vehicles and for the safety of commuters. Signalisation, which has a positive impact on safety, increases travel time and acts as a deterrent for experienced cyclists and a preference for inexperienced cyclists. While hills and road gradient, number of left turns, number of red lights and intersections, and traffic volume are the environmental factors that negatively affect the route choice of cyclists, the presence of various facilities on the road, scenery, greenery area, and blue space are listed as factors that positively affect the route choice. Land use and demography are also considered as other environmental factors affecting route choice. Although the number of studies centred on land use and cyclists' route choice is quite limited in the literature, residential areas negatively affect cyclists' route choice due to population density and some deterrent factors in cycling infrastructure, while mix, industrial and commercial land use positively affect cyclists' route choice. Table 1 shows various studies in the literature in order to better understand all these factors and their effects.

Table 1.; *List of various factors influencing cyclists' route choice*

Parameter	Factor	Influence on the route choice	Source
Travel Time	Traffic signalization	Negative	Band (2022), Lawrence & Oxley (2019), Lu et al. (2018)
	Cycle path	Positive	Band (2022)
	Number of intersections, left turns, and roundabouts	Negative	Lawrence & Oxley (2019), Lu et al. (2018), Beheshtitabar et al. (2014), Hood et al. (2013), Menghini et al. (2010)
	Traffic volume	Negative	Lawrence & Oxley (2019)
Traffic Safety	Lighting	Positive	Osama & Sayed (2017), Chen & Shen (2016), Reynolds et al. (2009), Wanvik (2009)
	Separated cycle lane	Positive	Trofimenko & Shashina (2021), DiGioia et al. (2017), Chen et al. (2016)
	Traffic signalization	Positive	Trofimenko & Shashina (2021), Lawson et al. (2013)
	Traffic volume	Negative	Lawrence & Oxley (2019)
	Painted cycle lane	Positive	Oršić et al. (2022), Schepers et al. (2011), Wall et al. (2016)

	Paved infrastructure	Positive	Dondi et al. (2011), Beheshtitabar et al. (2014)
Environmental Factors	Existing a facility on a route	Positive	Prato et al. (2018)
	Greenery area and blue space	Positive	Marquart et al. (2020), Prato et al. (2018)
	Gradient	Negative	Lu et al. (2018), Beheshtitabar et al. (2014), Broach et al. (2012), Scarf & Grehan (2005)
	Population density	Negative	Koch & Dugundji (2021), Prato et al. (2018)
	Residential area	Negative	Koch & Dugundji (2021), Prato et al. (2018)
	Commercial area	Positive	Prato et al. (2018), Winters et al. (2010)
	Industrial area	Positive	Prato et al. (2018)
	Mixed use area	Positive - Negative	Koch & Dugundji (2021) - Winters et al. (2010)

2.7. Research Gap and Aim of the Thesis

The factors that cyclists consider for route choice differ from those of drivers. For drivers, maximisation of utility comes to the forefront, while the most essential criterion for maximising utility is travel time (Teodorovic & Kikuchi, 1990). For cyclists, although travel time is important for route choice, safety, and environmental factors are also parameters that affect the final route decision (Ehrgott et al., 2012). Some of the results of various approaches on cyclists' route choice have been presented above. Although the studies generally reveal the factors affecting the route choice and how the preferences are affected, they do not provide an approach to integrate these factors together to develop a utility function and the probability with which route alternatives can be chosen. As shown in the studies of various researchers, fuzzy logic stands out as an alternative method that is both simple and more effective in obtaining the utility function compared to existing methods (Dhulipala et al., 2020; Dubey et al., 2013; Murat & Uludağ, 2008). Fuzzy logic has been used to evaluate route alternatives in an O-D pair for the route choice problem and to estimate the probability of choosing the available alternatives, but studies on generalising the route choice problem to a network have been limited. This study aims to develop a fuzzy logic model that can be used in small and medium-sized networks under limited data by considering the factors affecting the route choice of cyclists based on the route choice approach of fuzzy logic for an O-D pair in the literature. The developed fuzzy logic model will be tested using GPS data collected during the "fietstelweek" event in 2016.

3. METHODOLOGY

In this section, the fuzzy logic analysis model and its validation will be discussed. The flow diagram of the fuzzy logic model and model validation is shown in Figure 1. Accordingly, it is aimed to identify the factors affecting the route choice behaviour of cyclists through literature review. In this context, in the second step, the factors affecting route choice behaviour are discussed in detail under three headings: travel time, traffic safety, and environmental factors. The factors that form each of the three parameters are listed, and the effects of the factors on route selection are already presented in Table 1 above. Most of the factors in Table 1 will be used in the fuzzy logic model that is aimed to be created within the scope of this thesis. The main reason why some factors are not included in the analysis model is that Enschede, which is the study area, does not carry a distinctive function or the necessary data is not accessible/available due to the fact that cycling has an important place in daily life and the infrastructure has been shaped accordingly, unlike the cities and countries where studies in the literature are conducted in the Netherlands in general. After finding the necessary data sets and preparing the data, the analysis environment will be set up. Thus, the membership functions of the parameters considered will be determined, rule sets will be defined, and the appropriate fuzzy inference system will be selected for the model. The obtained results will be defuzzified. Thus, the attractiveness of all links in Enschede's cycling network for cyclists in terms of traffic safety and environmental factors will be obtained. While building the fuzzy logic model, GPS data collected during the "fietstelweek" event organised in 2016 is used. Since there is no systematic approach to be followed when building the fuzzy logic model, a threshold is set for finalising the rules to be used in the model. According to the threshold, the utility value of at least 80% of the routes followed must be close to or larger than the utility value of the fastest routes. The same threshold is valid not only for the fastest routes but also for the safest and most appealing route alternatives in the set of route alternatives. The fuzzy logic model thus constructed is applied to the city of Hengelo, which is similar to Enschede in terms of urbanization and socially and is located only 10 km west of Enschede. The fuzzy logic model is finalised if the utility values of the routes followed yield similar results in Hengelo. Then, the probabilities of route alternatives are found by using the logit formula, which has a simple formulation and is widely used to find the probabilities of being chosen of route alternatives according to their utilities. In addition to the logit formula, the C-logit formulation, which includes a commonality factor and reduces the probability of being chosen the routes if the route followed and a route alternative overlap, is used.

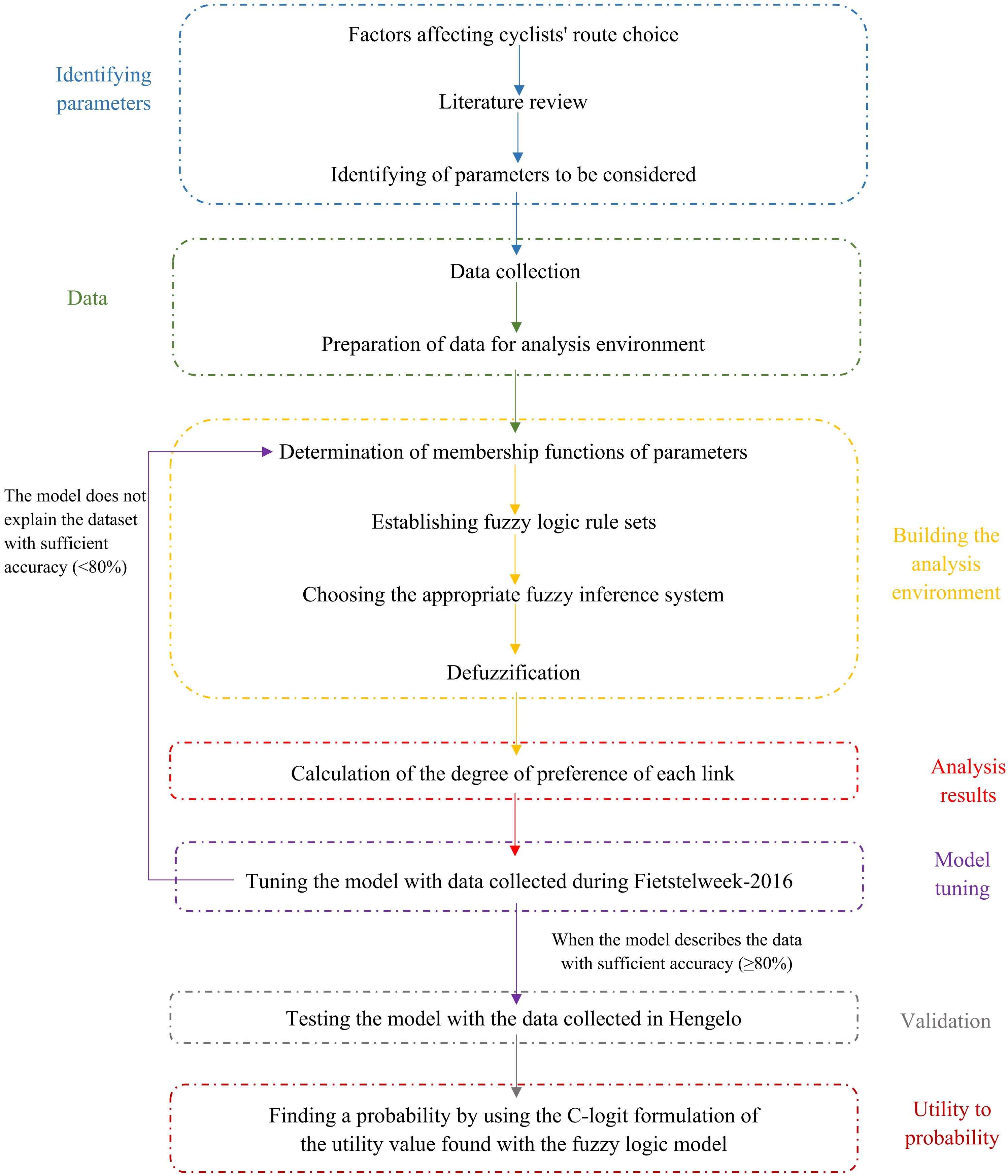

Figure 1.; *Methodological framework*

3.1. Study Area

The fuzzy logic model to be built within the scope of the thesis focuses only on the cycle trips that take place within the borders of Enschede within the event "Fietstelweek 2016". In this context, all links containing the cycle network within the borders of Enschede are matched with the data collected within the scope of "Fietstelweek 2016", and a total of 302 bicycle journeys taking place only within the borders of Enschede are filtered. Although cycling to other cities close to Enschede, such as Hengelo, Haaksbergen, Losser, Oldenzaal, etc. is frequently preferred, the study is limited to Enschede data due to the limitation of the analysis environment and the fact that the data sources to be used to build the analysis environment are open source and have various reliability problems. The location of Enschede on the map is shown in Figure 2.

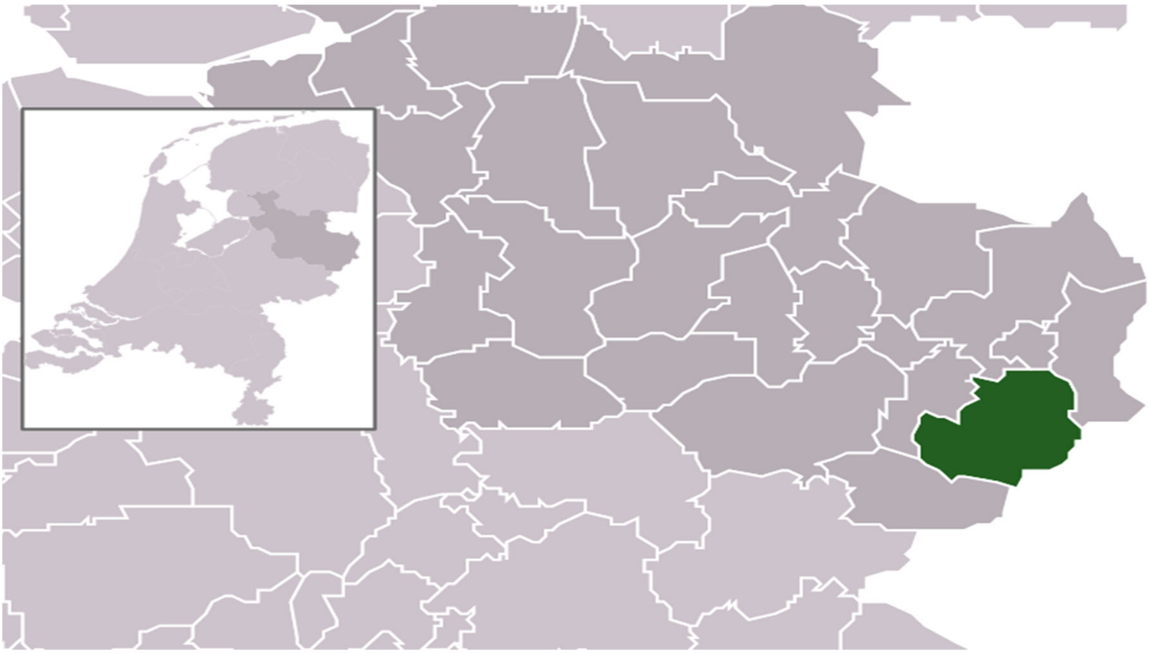

Figure 2.; *Location of the study area Enschede in the Netherlands*
(Source: Wikimedia Commons, 2009)

3.2. Identifying Parameters

The factors affecting the route choices of cyclists are discussed in detail in the second section in order to answer the first sub-question. Route choices are affected by travel time, traffic safety, and environmental factors. In the fuzzy logic model to be built within the scope of the thesis, traffic signalisation, link length, and traffic density on links that are not physically separated from motor vehicle traffic are considered as factors affecting travel time. Cyclists' safety and safer route choices are influenced by many factors. The Netherlands attaches importance to the construction and development of cycle infrastructure in its cities and shapes its infrastructure accordingly, both because of the widespread use of bicycles and because it follows a policy that encourages the increased use of bicycles (Government of the Netherlands, n.d.). In this context, with some exceptions, cycle lanes are marked in red, the cycle symbol indicating that it is a cycle lane is placed at intervals of 50-100 metres in residential areas and 500-750 metres outside residential areas, cycle lanes are physically separated from vehicle traffic on roads where the speed limit exceeds 50 km/h, where it is not possible to separate the cycle lane from vehicle traffic, the speed limit is limited to 30 km/h, and intersection points are controlled by signalisation to increase traffic safety (Bicycle Dutch, 2020). Therefore, paved infrastructure, traffic signalisation and painted cycle lanes will not be taken into account when assessing traffic safety on the links in the Enschede cycle network. In addition, road lighting is assumed to be given due consideration for cyclists, and lighting is another factor that will not be assessed for traffic safety, as lighting data on cycle lanes is not available. The assessment of traffic safety for cyclists considers whether the link is physically separated from motorised traffic, and if not, the traffic intensity to which the link is exposed. Environmental factors have a direct relationship with land use. High population density in residential areas has a negative effect on cyclists, while low population density in commercial or industrial areas has a positive effect on cyclists' route choice. In the analysis model, the land use in Enschede will be taken into account, and the effect of land use on the links passing near/along the land will be evaluated and included in the model. Since there is no significant elevation difference in Enschede, gradient is not taken into account. In addition, since a link-based analysis model has been built, the presence of a facility on the route is not taken into consideration in the environmental factors parameter in the analysis model.

3.3. Data

The data belonging to the parameters determined in order to build the analysis model are accessed by using different sources. All of the data to be used in the analysis model are obtained from secondary sources and are open-source data. Land use data in Enschede, the study area, is obtained from BBBike, an open-source map site. Infrastructure-related data, such as traffic lights, bicycle lanes, whether bicycle lanes are separated from motorised traffic by a physical barrier, and the type of road that forms the basis for traffic density are sourced using OpenStreetMap. For GPS data, as mentioned before, the data collected during the 'fietstelweek' event organised by the University of Breda in 2016 are used.

3.3.1. Fietstelweek GPS Data

The GPS data were collected as part of the Nationale Fietstelweek (National Bicycle Counting Week), a nationwide event from 19 to 25 September in which 42,658 people participated voluntarily (Fietstelweek, n.d.). The event was organised as a joint initiative between national agencies and companies to obtain more detailed and reliable information about the cycling behaviour of cyclists in the Netherlands, and to plan and make improvements based on the information collected. GPS data was collected using the Fietsel-app, developed specifically for the event, and participants were asked to answer a questionnaire with sociodemographic and travel behaviour questions in order to identify the GPS data (Fietstelweek 2016, n.d.). Within the scope of the event, a total of 416,376 cycle journeys covered 1,786,147 kilometres across the country. The average speed of cyclists during these journeys was 15.8 km/h (Fietstelweek 2016, n.d.). The sociodemographic and travel behaviour data collected during the event is not available as of 2023, the year the thesis was written. Therefore, the characteristics of the trips made, and the purpose of the trips are not known.

The GPS data collected during the event is available in csv and shp file format. The shp file format shows the location of the GPS data collected during the event on the cycling network in the Netherlands. The csv file contains the unique participant number of each participant, the name of the links that the cyclists travelled through in the shp file, the average speed of the cyclist on the link, whether the cyclist travelled on weekdays or weekends and the time of the trip. The time information includes only the hour the cyclist started travelling and does not include minutes or seconds.

In Enschede, which is chosen as the study area, a total of 302 cyclists participated in the event. Participants travelled approximately 1750 km within the borders of Enschede. Figure 3 shows the GPS data collected from the participants during the event within the borders of Enschede. The average speed of the participants was 15.68 km/h.

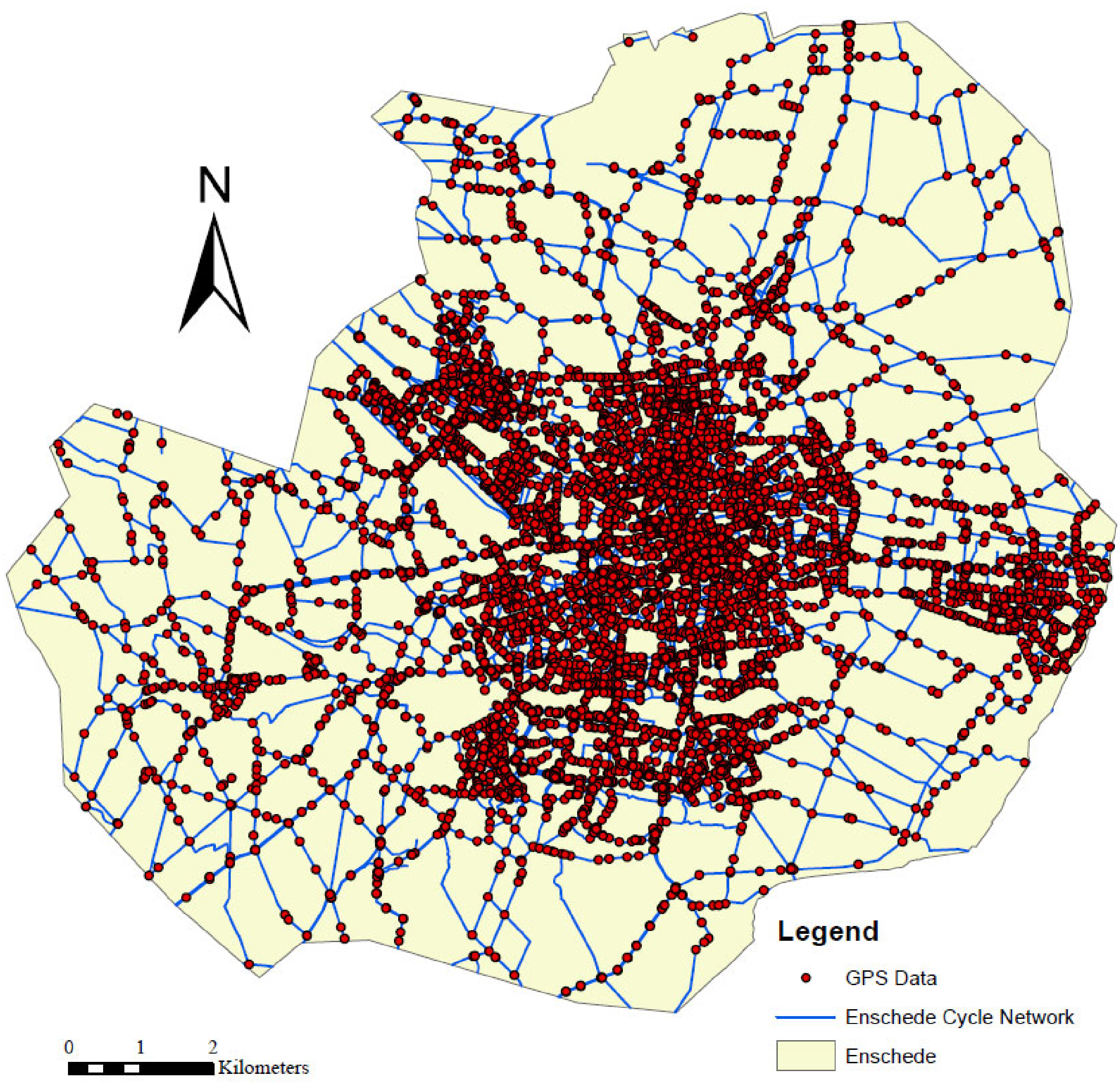

Figure 3.; *GPS data of the participants of Fietstelweek 2016 in Enschede*

According to the data, data was collected only once from all participants in Enschede. In other words, each participant attended the event for only one day. During the week of the event, 294 of the journeys were made on weekdays and 8 on weekends. Most trips were made on weekdays between 8:00-09:00, 16:00-17:00 and 17:00-18:00. Approximately 47% of the journeys took place in these three time periods. Figure 4 shows the distribution of journeys in Enschede during the event by weekday and weekend hours.

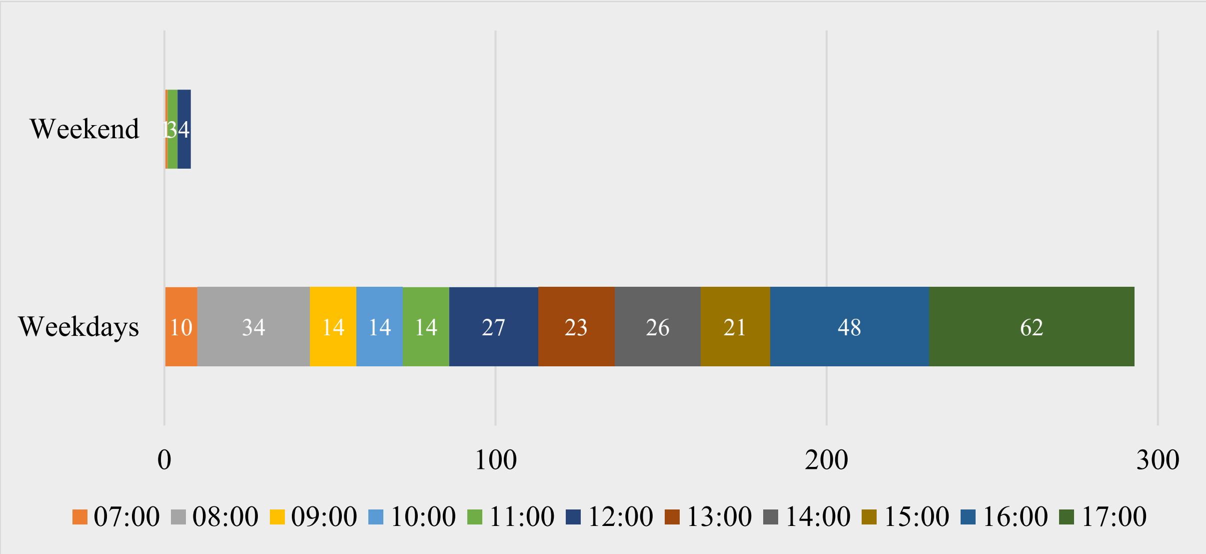

Figure 4.; *Distribution of cycle trips within the scope of the Fietstelweek by weekdays/weekends and hours*

The distribution of the trip lengths of the cyclists participating in the event within the borders of Enschede is shown in Figure 5. The average distance travelled per cyclist is 5.79 km. 69% of the cyclists travelled 7 km or less. While 39% of the journeys are between 1-4 km, 26% of the trips are between 4-7 km. The fewest trips by distance travelled are under 1 km, with a total of 13 trips under 1 km. Approximately 17% of cyclists travelled over 10 km. 42 cyclists travelled between 7-10 km, which is about one seventh of the total trips.

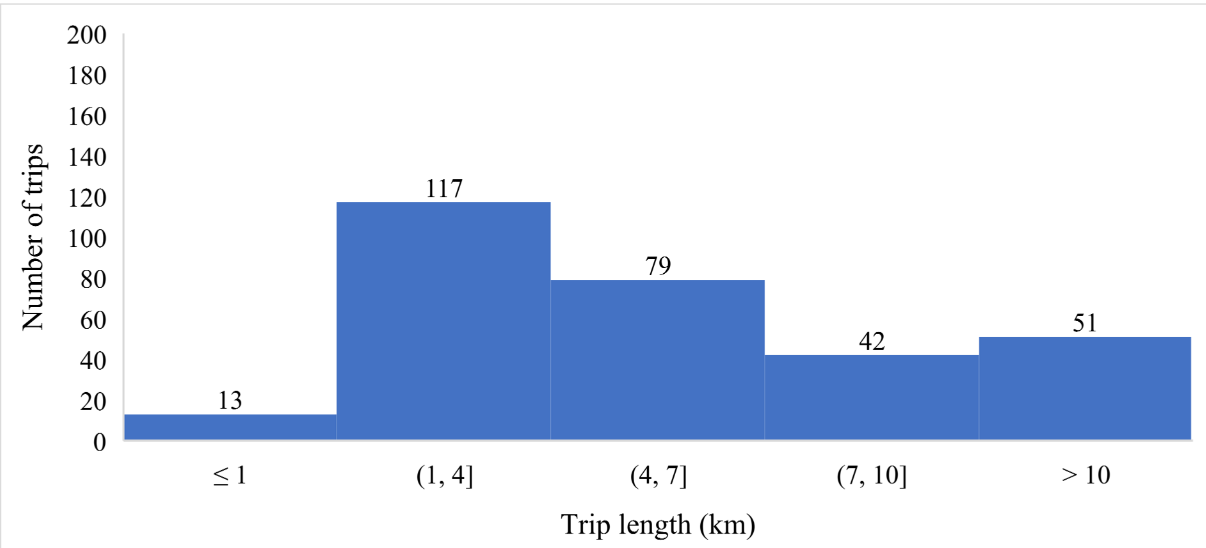

Figure 5.; *Distribution of cycling trips during the Fietstelweek by distance travelled*

Figure 6 shows the routes followed by the participants within Enschede. The blue colour indicates the cycling network in Enschede, while the other colours on the cycling network indicate the route followed by each cyclist.

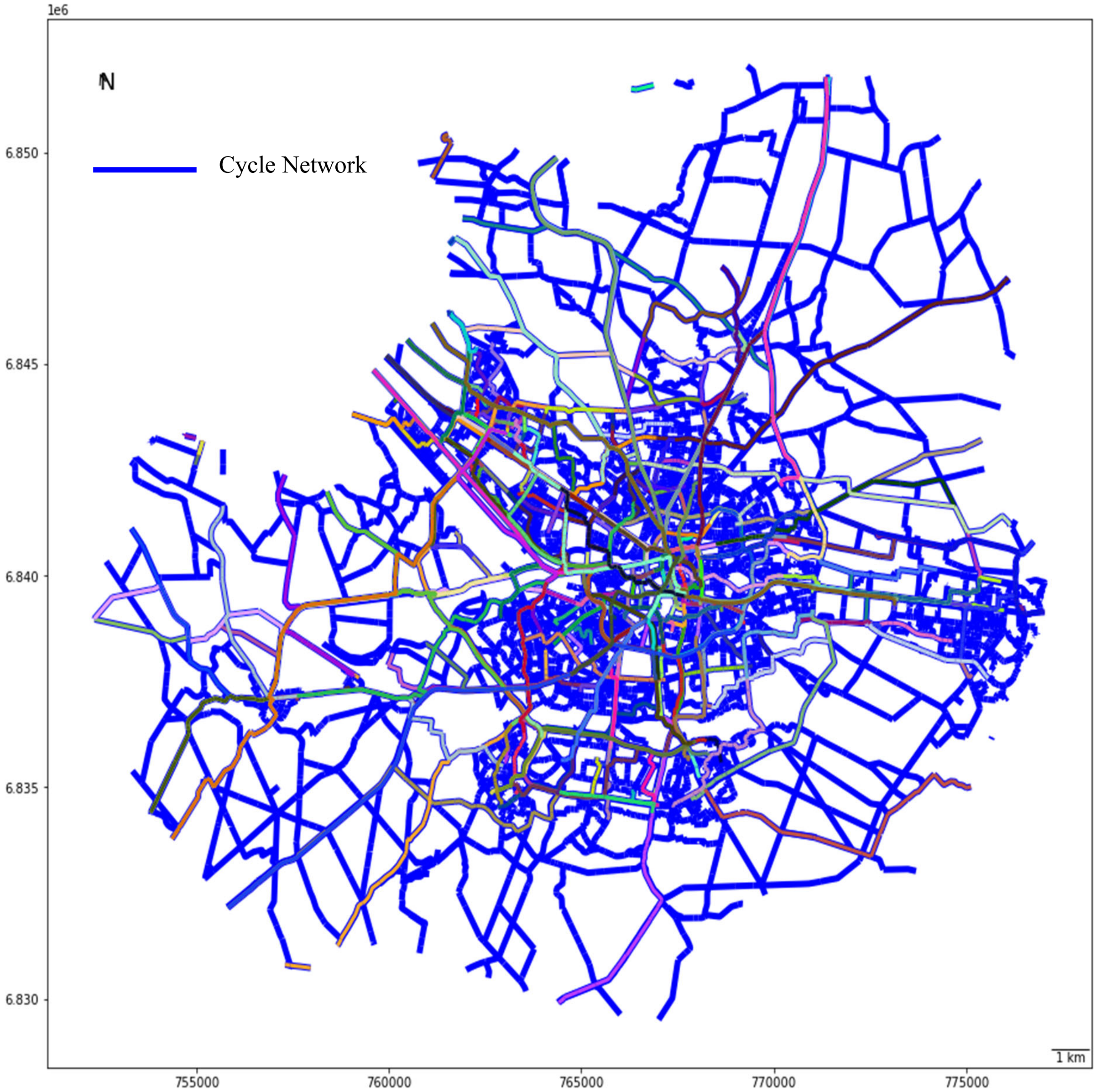

Figure 6.; *Routes followed by participants on the cycling network in Enschede*

3.3.2. Land Use

Land use data for the study area Enschede is obtained from BBBike. BBBike is a free open source map site that allows saving data on Planet.osm in various formats such as OSM, PBF, SHP, GeoSJON. Users provide data access by selecting the format they want and the area they want to work on. The maximum area that can be accessed at one time through the site is limited to 24 million square kilometres or 500 MB file size.

In the land use map downloaded from BBBike, 30 different land use types are grouped under eight different labels by grouping similar land uses. The land use in Enschede is shown in Figure 7.

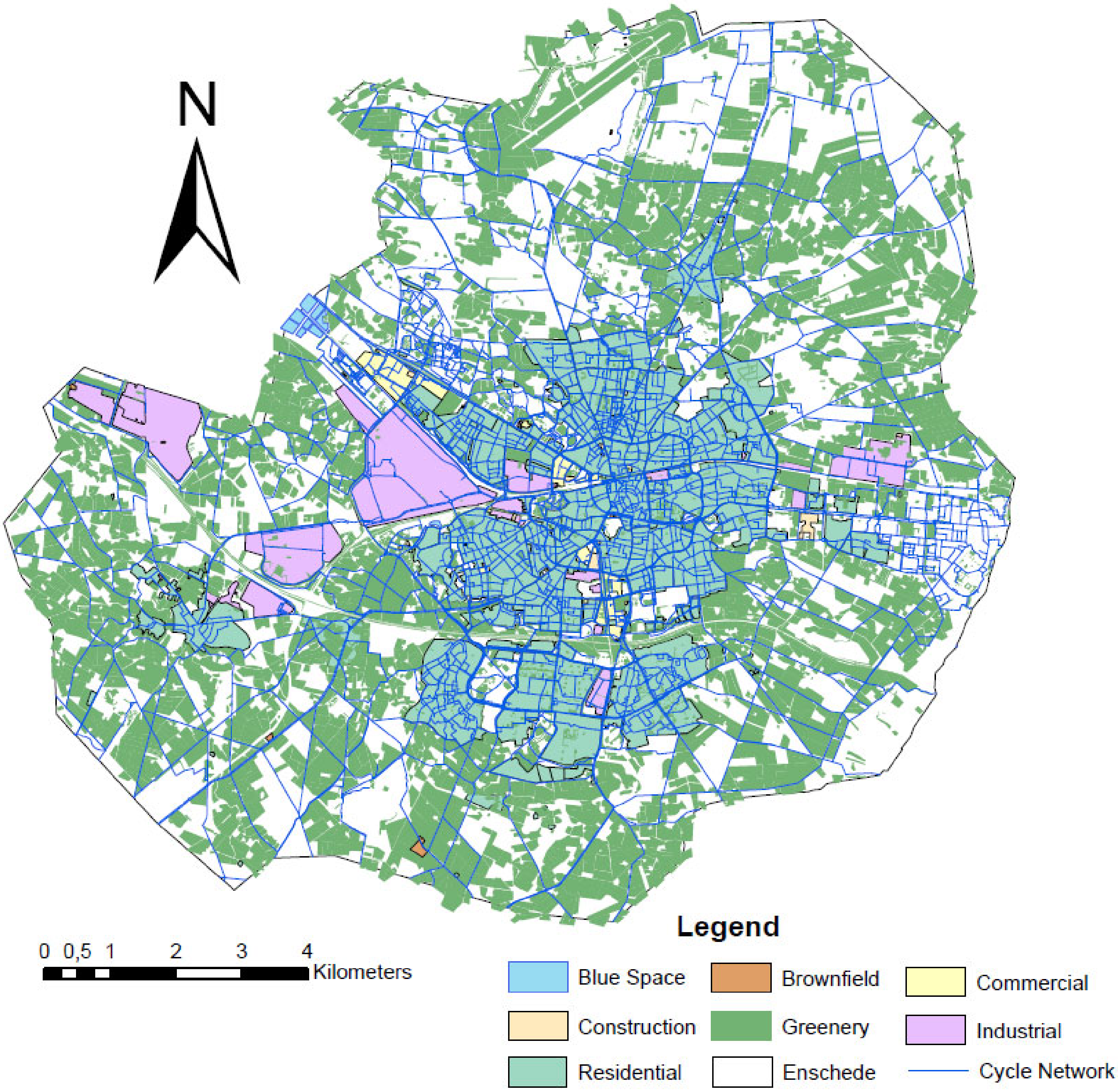

Figure 7.; *Land use in Enschede*

Enschede has a total area of 142.72 km^2. Within the total area, Brownfield is the land use label with the smallest land use among the labels with an area of approximately 0.17 km^2. Land use labels with Construction and Blue Space labels have very close areas to each other. These labels have an area of 0.21 km^2 and 0.22 km^2, respectively. The total area of land use with these three labels is 0.42% of Enschede's area. The size of commercial areas in the city is approximately 1.18 km^2. The total size of industrial areas, the majority of which are located east of the city centre, is measured at 6.60 km^2. The sum of the commercial and industrial areas in Enschede accounts for 5.45% of the total area. The total area of residential areas clustered around a radius of about 3 km around the city centre is 22.70 km^2, corresponding to 15.91% of the total area. Finally, due to the large amount of forested areas in the city, the total area of green areas, which are regularly spread around the city and cover the largest area, is 58 km^2. Green areas constitute approximately 41% of the total city area. The 53.64 km^2 area, which is 38% of the total area, is not classified under any land use label.

3.3.3. Cycle Network of Enschede

The cycling network of Enschede is obtained as part of the Fietstelweek data. The data shows the GPS data collected during the event and the locations of these GPS data on the cycling network in the Netherlands. The event data does not contain detailed information about the links that constitute the bicycle network. For this reason, OpenStreetMap (OSM) is used in order to use the bicycle network obtained within the scope of the activity more effectively. Map data, such as bicycle path, traffic lights, whether the bicycle path is physically separated from motor vehicle traffic, etc., are thus included in the existing bicycle network. OSM is a free open-source geographical database that is created and updated with the co-operation of volunteers from all over the world. Since the maps in OSM are produced as a result of the collaborative work of volunteers and no commercial concerns are pursued, there may be errors in both the characterisation of the data and the validation. Therefore, it should be kept in mind that the data obtained through OSM may not be accurate.

The cycle network in Enschede can be characterised in three ways: roads without a separate or marked cycle lane, roads with a separate and marked cycle lane but not separated from motor vehicle traffic, and roads dedicated to cyclists, separated from motor vehicle traffic by a physical barrier. The cycling network in Enschede and the characteristics of the links that constitute the network are shown in Figure 8.

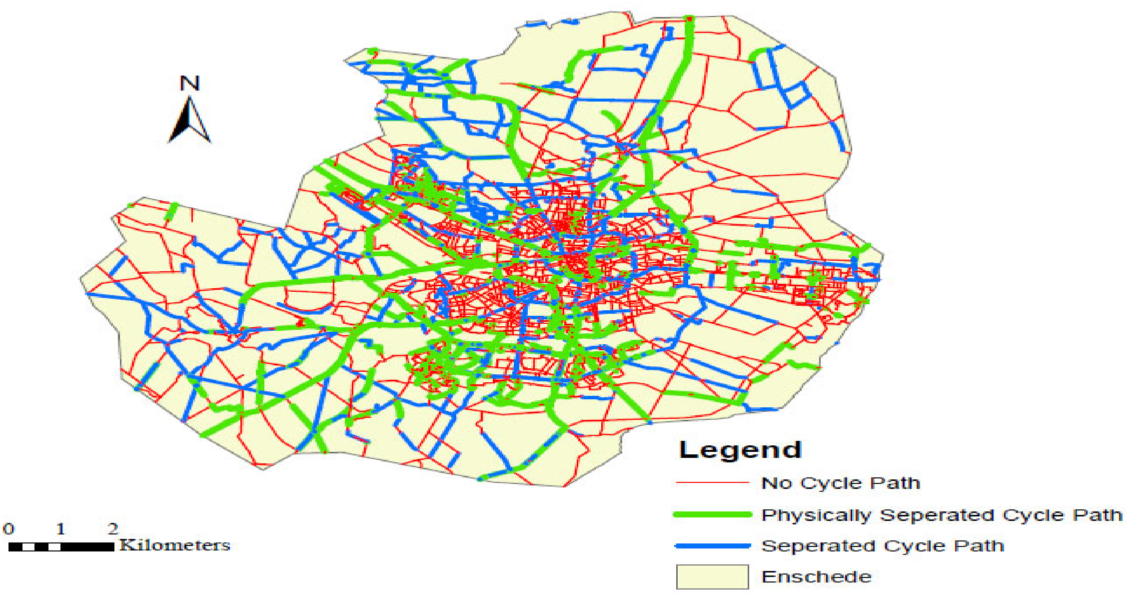

Figure 8.; *Characteristics of the links that form the Enschede bicycle network (based on data from OSM)*

Enschede has a total cycle route network of 1522 km, of which 843 km, or 55%, do not have a separate cycle lane. The total length of separated roads is 679 km in the city, where there is a separate road for cyclists on roads with heavy motorised traffic. Of the separated roads, 46 % are physically separated from motorised traffic, while 367 km are separated from motorised traffic only by means of cycle lane markings.

3.4. Building the Analysis Environment

Information about building a fuzzy logic model from data obtained from open sources, determining the membership functions of the parameters considered, associating the parameters with each other with rule sets and defuzzification will be given both in the context of this problem and the theoretical framework of fuzzy logic.

Among the travel time, environmental factors, and traffic safety parameters to be used in the fuzzy logic model, travel time is directly associated with the route chosen by the road user. Therefore, no result is going to be obtained on a link basis. Environmental factors and traffic safety parameters will be evaluated by considering the factors previously discussed within the scope of the effect of different factors on the route choice behaviour of cyclists, and traffic safety and environmental factor scores will be assigned to each link. Thus, it will be possible to analyse any desired route for any O-D pair on the network in terms of traffic safety, environmental factors, and travel time with a fuzzy logic model. If different routes are to be compared for the same O-D pair, the fuzzy logic utility values obtained for the routes will be compared and an approach will be developed to determine which route will be chosen with which probability.

3.4.1. Fuzzy Logic

3.4.1.1. Fuzzy Sets

In classical logic, a set is defined as a collection of certain distinguishable objects. According to this definition, it is known exactly to which set any element belongs. In fuzzy logic, the elements of a set can be elements of one or more clusters according to their membership degrees, which indicate the degree to which an element belongs to a cluster. In other words, for any element x of a fuzzy set, the expression 'element x belongs to set A with membership degree $\mu_{\tilde{A}}(x)$' is used instead of 'belongs to set A' or 'does not belong to set A'. The aim of the theory of fuzzy sets is to assign a degree of membership to vague, ambiguous, and difficult to define concepts in order to define, specify, and express them more easily (Ross, 2009). Specificity comes from the transformation of the theory of two-valued sets into the theory of multi-valued sets (Ross, 2009).

Operations on fuzzy sets

The most significant identifier of a fuzzy set is its membership function (Ross, 2009). Therefore, basic set theory operations are defined through the membership functions of fuzzy sets. The basic set theory operations proposed by Zadeh, where $\tilde{A}$ and $\tilde{B}$ are fuzzy sets defined on the universal set X, are shown below.

Equality: The necessary and sufficient condition for the fuzzy sets Ã and $\tilde{B}$ to be equal is that the degrees of membership for all points of the fuzzy sets defined in the universal set X are equal. Accordingly, if $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$ (x) for $\forall x \in X$, then Ã = $\tilde{B}$.

Intersection (AND): Let X be a universal set, and the intersection of fuzzy sets Ã and $\tilde{B}$ is defined by the membership function $\mu_{\tilde{A} \cap \tilde{B}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$ for $\forall x \in X$ and is shown in Figure 9.

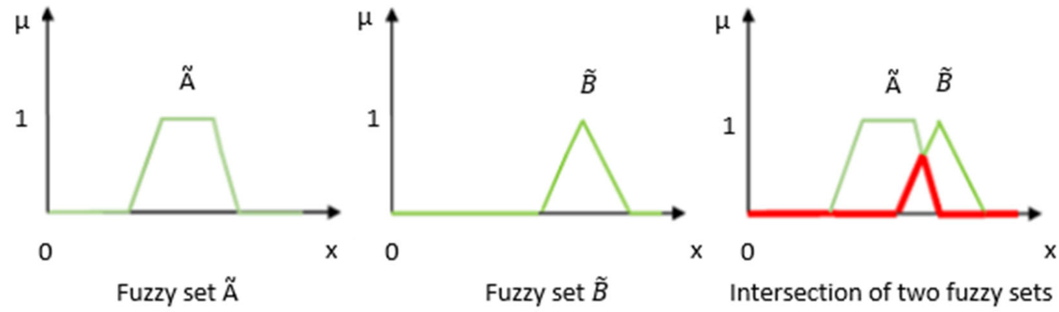

Figure 9.; *Intersection of two fuzzy sets (Source: Tutorialspoint)*

Union (OR): The union of fuzzy sets Ã and $\tilde{B}$, where X is a universal set, is defined by the membership function $\mu_{\tilde{A} \cup \tilde{B}}(x) = \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$ for $\forall x \in X$ and is shown in Figure 10.

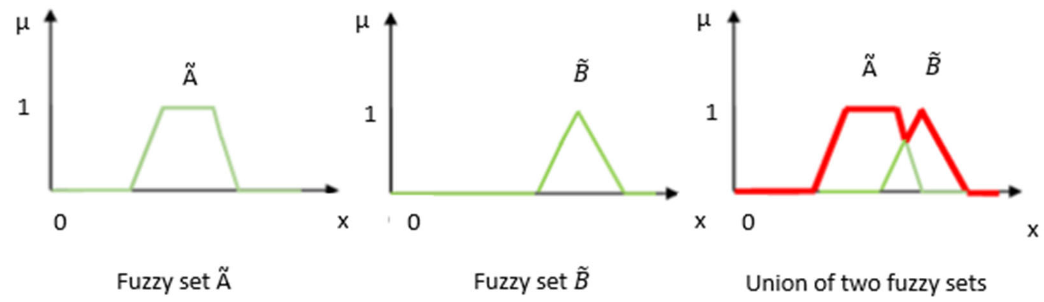

Figure 10.; *Union of two fuzzy sets (Source: Tutorialspoint)*

Complement (NOT): The complement of a fuzzy set Ã, where X is a universal set, is denoted by $\bar{\tilde{A}}$ and is defined by the membership function $\mu_{\bar{\tilde{A}}}(x) = 1 - \mu_{\tilde{A}}(x)$ for $\forall x \in X$.

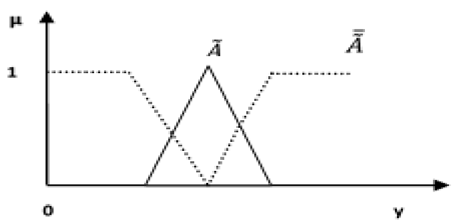

Figure 11.; *Complement of a fuzzy set*

3.4.1.2. Membership Function

The membership functions required for fuzzy logic operations are a group of expressions consisting of linguistic qualifiers. A numerical range represented by a fuzzy word or expression under consideration can be determined by analysts who have knowledge of that expression. For example, the range of variation of the temperature in Enschede can be stated to be from -10 ºC to 30 ºC. This range specifies the range in which the elements of the temperature set for Enschede can be placed. Thus, the whole temperature space is determined. However, in colloquial language, the temperature space can also be thought of as consisting of several subsets, such as very hot, hot, warm, cold, very cold. If it is necessary to decide where the range of each subset begins and ends, it can be stated that each of these subsets is not overlapping, but as if they are the continuation of each other at the border. There is no overlap in the range estimates made here. It may not always be correct to conclude that these intervals are formed by precise boundaries without overlapping each other.

Another question is whether the temperature degrees belonging to each subset are of the same importance. For example, as we approach the lower and upper ends of the warm range, we can expect transitions towards its neighbouring sub-sets of warm at the bottom and cold at the top, so it cannot be said that the intervals coinciding with the transition zones will have the warm property in full. Thus, it can be concluded that the temperatures falling in each subset will lose their relative importance near the extremes of that subset. From this point of view, if a value called the degree of importance is considered in a subset, it can be said that the highest values of this value will be in the middle and the lowest values will be at the lower and upper ends. The mathematical representation of this mental experiment is given in Figure 12.

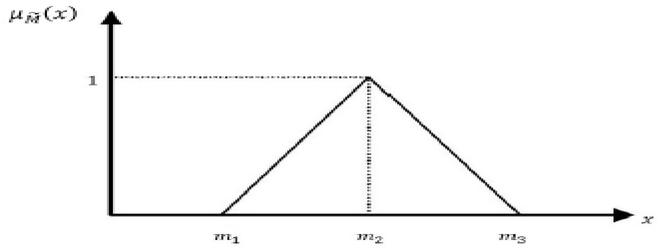

Figure 12.; *A membership function*

3.4.1.3. Fuzzy Logic Rules and Inference

In order to build a fuzzy logic model, certain rules must be defined, and an inference must be derived from this set of rules. There are two approaches or connectives that are frequently used to combine a large number of rules in a fuzzy logic model. The first one is to combine all the rules defined in the fuzzy logic system with the conjunction "AND". The intersections of the rules are taken with the equivalent of this conjunction in set theory. The second approach is the use of the conjunction "OR", and an inference is drawn by taking the intersections of independent variables with each other.

In the fuzzy logic model, independent variables are connected to each other with "IF-THEN" structure. In the model, the part between 'if' and 'then' is called precondition, and the part after then is called inference. In the rule-based fuzzy logic model, preconditions and inference are obtained by fuzzification separately. Mamdani fuzzy-inference system developed by Mamdani (1976) is widely used to obtain a result by combining fuzzy rule-based expressions with 'and/or' connectives.

3.4.1.4. Defuzzification

As the inputs of a fuzzy logic model are fuzzy, its outputs are also fuzzy. However, it is not desirable that the values to be used in decision making are not fuzzy; therefore, fuzzy values should be converted into a classical value. The transformation of a fuzzy value generated as a result of a fuzzy logic model into a number value used in classical mathematics is referred to as "defuzzification" (Ross, 2009). Defuzzification is performed by means of the membership functions of the fuzzy set obtained as a result of fuzzy operations. The classical value obtained as a result of the defuzzification is a value between the left and right components of the fuzzy set (Ross, 2009). There are a wide variety of methods in the literature. Only centroid method is covered in this study, since it is used as the defuzzification method within the scope of this thesis.

Centroid Method

It is a method in which all elements in the fuzzy set affect the result with their membership values. Centroid Method, where the membership function of an element in the fuzzy set is μ(z),

$$z^* = \frac{\int \mu(z) z dz}{\int \mu(z) dz} \tag{3.1}$$

is expressed in the form. Since it is algebraically more efficient, the centroid method is one of the most widely used defuzzification methods (Ross, 2009).

3.4.2. Traffic Safety and Environmental Factors

Traffic safety and environmental factors are assessed on the basis of each link forming the cycle network. When evaluating the traffic safety of each link, membership functions are generated by considering whether the cycle path is separated from motorised vehicle traffic and which road classification (pedestrian, primary, secondary, etc.) the link belongs to. Similarly, when evaluating a link in terms of environmental factors, membership functions are generated by evaluating the land use around the link. In the literature review studies, it has been concluded that the presence of blue/green areas increases the attractiveness of a route, industrial and commercial land use has a positive contribution to route choice, while residential and construction areas have a negative effect on route choice. This framework is used in the fuzzy logic model to be constructed. Membership function values and rule sets are also constructed by considering this framework.

A map is prepared by using the scores obtained as a result of analysing the bicycle network with the fuzzy logic model in terms of environmental factors and traffic safety parameters. The map can be used as a guide for road users to determine a safe and attractive route. The steps in Figure 13 are followed to get and visualise the attractiveness values of the links forming the bicycle network in Enschede by using the environmental factors and traffic safety parameters.

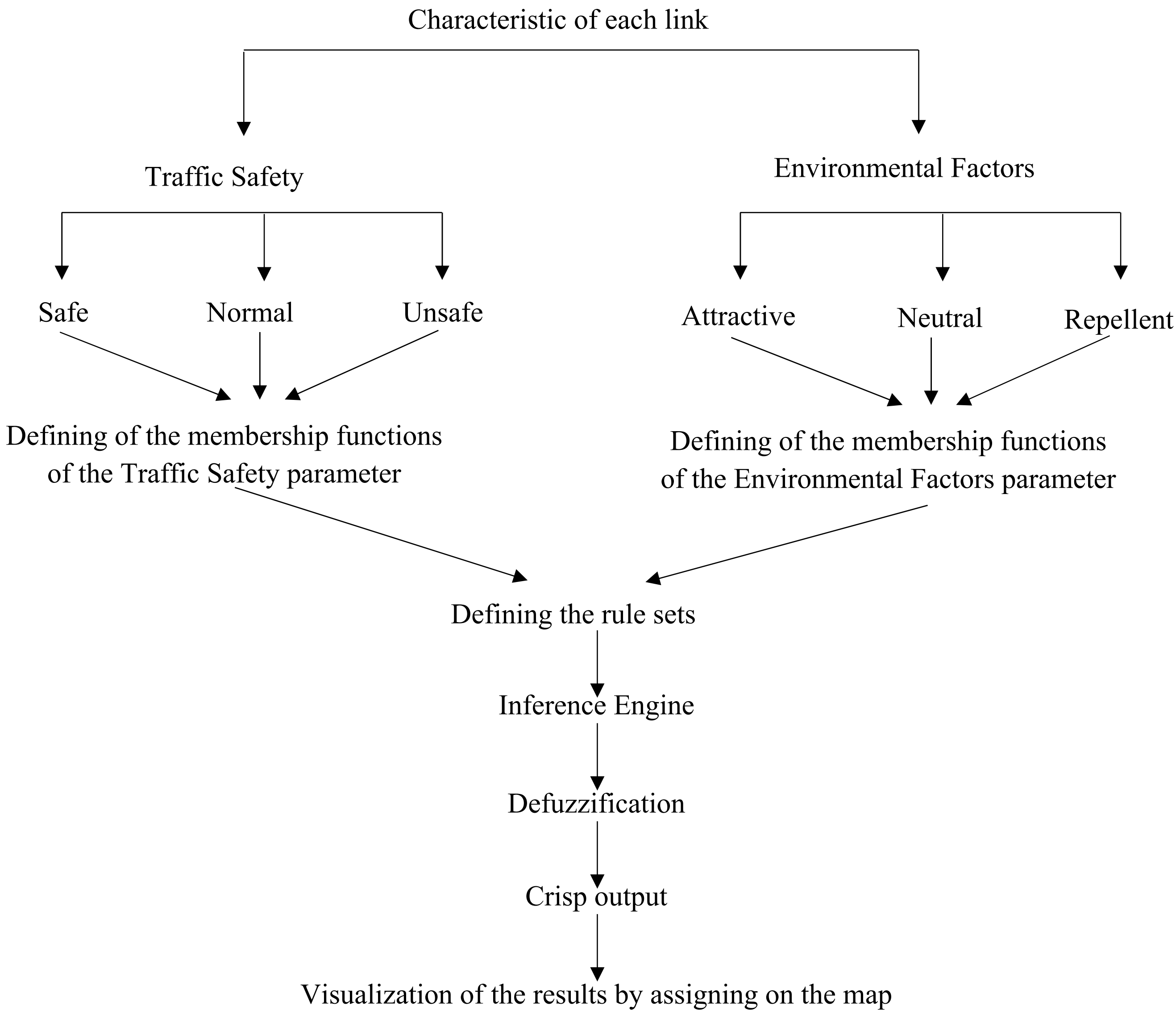

Figure 13.; *Evaluation and visualisation of links in the frame of traffic safety and environmental factors with fuzzy logic model*

In the traffic safety parameter, the links are grouped in three different ways as safe, normal, and unsafe. In the traffic safety parameter, each link takes a value between 0 and 60 considering whether it is separated from the motorised vehicle traffic on the road segment on which it is located, if separated, the separation is physically or not, and the characteristics of the road segment for roads that are not physically separated. For the Traffic Safety parameter, it is a totally arbitrary choice for the values to vary between 0-60. One of the most basic features of fuzzy logic is that the membership function for a parameter can be defined within the desired value range. Thus, comparable values can be obtained due to the attribute it has while calculating the utility value. Values between 0-30 are characterised as 'Unsafe', values between 15-45 as 'Normal', and values between 30-60 as 'Safe' and represented by triangular fuzzy numbers. The triangular fuzzy number function showing traffic safety is shown in Figure 14.

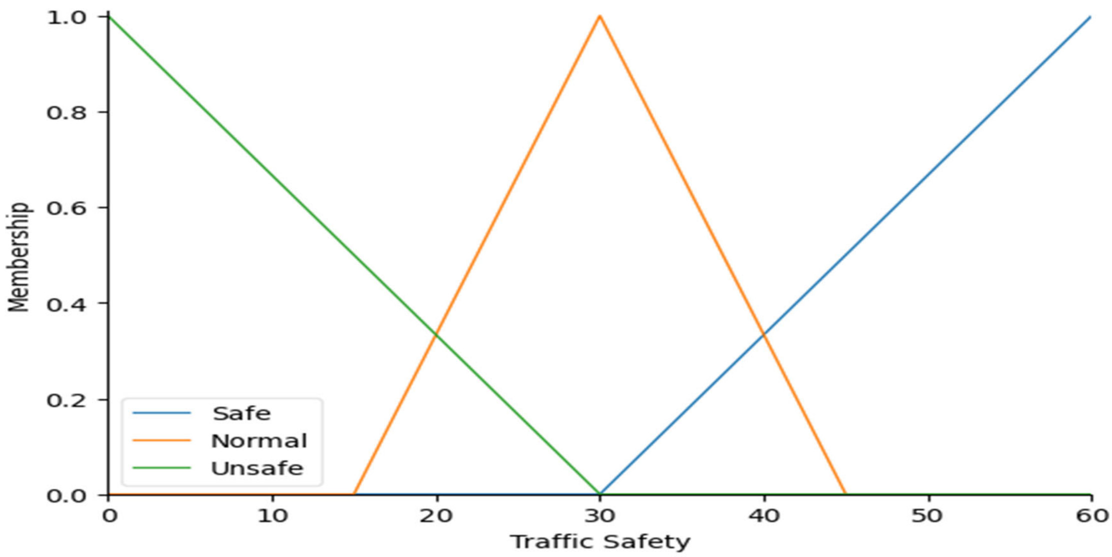

Figure 14.; *The triangular fuzzy number function used for the Traffic Safety parameter*

In the bicycle network from open sources, the links are grouped under five headings as pedestrian, cycleway, primary, secondary, and tertiary, and some links are not included in any group and are excluded from the classification. In this context, roads physically separated from vehicular traffic are characterised as 'Safe' and assigned to the 'Safe' fuzzy set with a high membership rate. All links belonging to this subset are assigned scores ranging from 50 to 60 and membership degrees are generated. A wide range of values are assigned to cycle paths that are separated from motorised traffic by markings rather than physical barriers. These values ranged from 20 to 50 depending on the characteristics of the road segment. For example, if the link is on a primary road and separated from vehicular traffic by markings instead of physical barriers, a range of values between 20 and 40, in other words safe-normal-unsafe, is assigned. This is based on the assumption that such roads can be categorised as safe, unsafe, or normal depending on the experience of the road user. Finally, links that are not separated from vehicular traffic are assigned values ranging from 0 to 30 depending on their nature on the road. No link in this group has taken a value higher than 30, where the 'Safe' fuzzy set starts. Therefore, a link that is not physically or markedly separated does not have a degree of membership in the 'Safe' fuzzy set. Table 2 shows the range of values assigned to the links that make up the cycling network according to the nature of the cycle path and the grouping of neighbouring vehicular/pedestrian traffic.

Table 2.; *Traffic safety fuzzy number values assigned according to link characteristics*

Link characteristics	**Physically Separated**	**Separated**	**Unseparated**
Pedestrian	55/60	35/50	20/30
Cycleway	55/60	40/50	20/30
Primary	50/60	20/40	0/20
Secondary	50/60	30/45	5/20
Tertiary	55/60	35/45	10/30
Unclassified	50/60	20/40	10/30

Similar to the traffic safety parameter, the environmental factors parameter is grouped into three different fuzzy sets as repellent, neutral, and attractive to indicate the attractiveness of the link. The attractiveness of a road user's trip on a link is associated with the land use to the left and right of the link. Depending on how the land use on both sides of a link is used, a link is assigned a value between 0 and 50 in the environmental factors parameter. In the triangular fuzzy number function, the environmental factors value is defined as 'Repellent' if it is between 0-25, 'Neutral' if it is between 15-35, and 'Attractive' if it is between 25-50. The triangular fuzzy number function of the environmental factors parameter is shown in Figure 15.

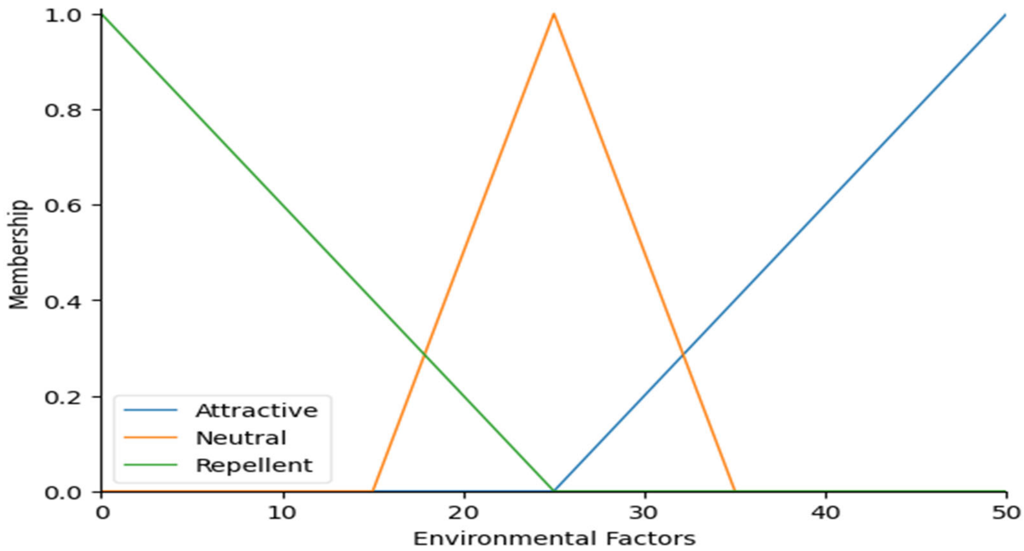

Figure 15.; *The triangular fuzzy number function used for the Environmental Factors parameter*

Using the land use data obtained from BBBike, each link is assigned a value according to the surrounding land use. The links according to the surrounding land use are found by using the selection by location feature in ArcGIS software. Land use is grouped into seven categories as Blue Space/Greenery, Brownfield, Commercial, Construction, Industrial, and Residential. Links with Blue Space/Greenery land use on both sides are considered as the links with the highest membership to the 'Attractive' fuzzy set, and values ranging between 45-50 are assigned to links with this attribute. For links with Blue Space/Greenery on one side and other land use labels on the other side, value ranges are assigned considering the effect of land use on attractiveness. Links neighbouring areas with Brownfield, Construction, and Residential land use labels are considered unattractive. If there are areas with unattractive land use on both sides of a link, these links are assigned values ranging from 0 to 20. Thus, the membership degree of these links to the 'Attractive' fuzzy set is 0. Values ranging from 0 to 20 are defined by considering that in areas with unattractive land use, some road users may consider these links as unattractive, and some road users may consider them as neutral. Therefore, they are defined in a wide range. The links adjacent to Commercial and Industrial areas are also defined as having a quality that increases attractiveness based on the literature. However, it is assumed that these two land use labels do not increase attractiveness as much as Blue Space/Greenery. Table 3 shows the fuzzy values assigned to the links depending on the land uses on both sides of a link.

Table 3.; *Environmental factors fuzzy number values assigned according to the land use*

Land Use	**Blue Space / Greenery**	**Brownfield**	**Commercial**	**Construction**	**Industrial**	**Residential**
Blue Space / Greenery	45/50	35/40	40/45	30/35	40/45	30/35
Brownfield	35/40	10/20	25/35	10/20	25/35	5/15
Commercial	40/45	25/35	30/40	15/25	30/40	15/25
Construction	30/35	10/20	15/25	0/15	15/25	0/15
Industrial	40/45	25/35	30/40	15/25	30/40	15/25
Residential	30/35	5/15	15/25	0/15	15/25	0/15

3.4.3. Travel Time

While calculating the travel times of the links, several assumptions have been made for the travel speed by considering traffic lights, traffic density, pedestrian density in pedestrian-dominated areas, and rush hour factors in addition to the link length. The travel time on each link is calculated as the ratio of the link length to the speed on the link. The travel speed of a cyclist on a link is assumed to be 15.68 km/h, which is the average travel speed of the cyclists participating in the event in Enschede. Since the travel time is affected by traffic density and pedestrian density, two different travel times are used for vehicle and pedestrian peak hours. One of the most significant factors affecting the travel time is the traffic lights. Since the duration of the traffic lights in Enschede changes dynamically depending on the traffic density, it is assumed that the waiting time at the traffic lights increases due to the increased traffic density between 07:00-09:00 and 16:00-18:00 during rush hour. Outside of rush hour, it is assumed that a cyclist waits for 25 seconds when he/she encounters a traffic light and loses 23 seconds at a traffic light, assuming that he/she needs 8 seconds to reach his/her average speed again when the light turns green. In the rush hour, due to the increased traffic density, the waiting time increases by 60% to 40 seconds, and it is thought that the cyclist loses a total of 48 seconds to encounter a traffic light. According to OSM data, the time lost at traffic lights has been added to the travel time on links with a total of 142 traffic lights in 31 different locations. Figure 16 shows the locations of the traffic lights in Enschede based on OSM data. In addition, on primary roads that are not physically separated from vehicle traffic by a physical barrier, the speed of the cyclist in front is affected by the cyclist travelling in front of him due to increased cycle traffic during rush hours. Overtaking the cyclist in front may not always be a safe and rational move due to heavy vehicle traffic. For this reason, if the roads characterised as primary and secondary roads in rush hour are not separated from vehicle traffic by a physical obstacle, the travel times of cyclists on these links are increased by 15% and 10%, respectively, taking into account the impact on the speed of cyclists. Finally, it is assumed that cyclists will reduce their speed by 20% in order to ensure their own safety and not to endanger the safety of pedestrians due to the increased pedestrian traffic between 17:00-20:00 on weekdays and after 13:00 on weekends in the city centre where pedestrian traffic is intense.

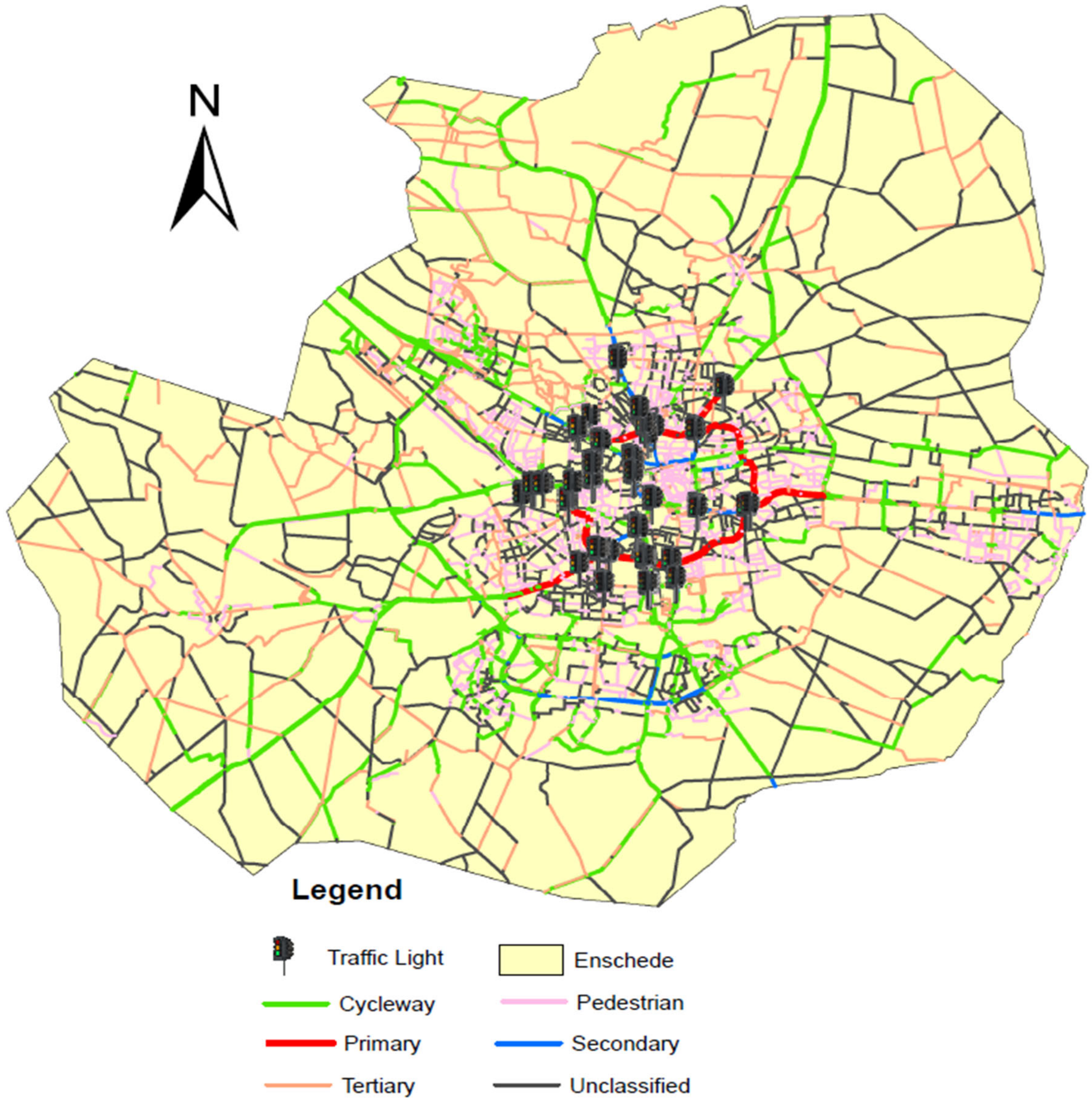

Figure 16.; *Characteristics of the links in the Enschede cycle path network*

It is not possible to evaluate the fuzzy logic membership functions for travel time only on link basis in terms of building a fuzzy logic model. Therefore, the travel time is determined by considering the entire route followed. Based on the shortest possible route from origin to destination, the travelling time of the route followed can be expressed as fuzzy. In this context, the 'k shortest path' algorithm is used to find the shortest path in terms of travelling time for any O-D pair to be determined and other shortest path alternatives if necessary. Depending on how different the route alternatives are from the shortest path in terms of percentage, fuzzy number values are determined, and membership functions are constructed. Five different fuzzy sets are defined as 'very short, short, normal, long, very long' depending on the percentage difference of the travel time from the shortest path. The main purpose of defining the travel time with five different fuzzy sets different from traffic safety and environmental factors is to minimise the loss of information. Since travel time is the most significant factor in the choice of a route as shown by various studies, five fuzzy sets are defined for travel time instead of three in order to obtain more precise values. In order for the defined fuzzy sets to take a value between 0 and 100, another fuzzy set is defined, and road alternatives are given a value. Figure 17 shows the quantification of the Travel Time parameter of the travel time on a route compared to the shortest travel time between route alternatives. The figure on the left shows the membership degrees in which fuzzy sets a route is classified depending on the percentage difference between the travel time of a route and the travel time of the shortest route between the same O-D pair. The figure on the right shows the value between 0-100 for the Travel Time parameter according to the membership degrees in which it is classified. Accordingly, if the travel time is up to 10% longer than the shortest route, it is defined to belong to the 'very short' fuzzy set with a triangular changing membership degree and it is provided to get a score ranging between 75 and 100 depending on the membership degree. The 'very long' fuzzy set is defined as belonging to the travel time fuzzy set with varying degrees of membership starting from 30% up to 60% travel time difference. If the travel time difference is 60% and above, the travel time belongs to the 'very long' fuzzy set. A travel time belonging to the 'very long' fuzzy set takes a value between 0 and 40 depending on the membership degree.

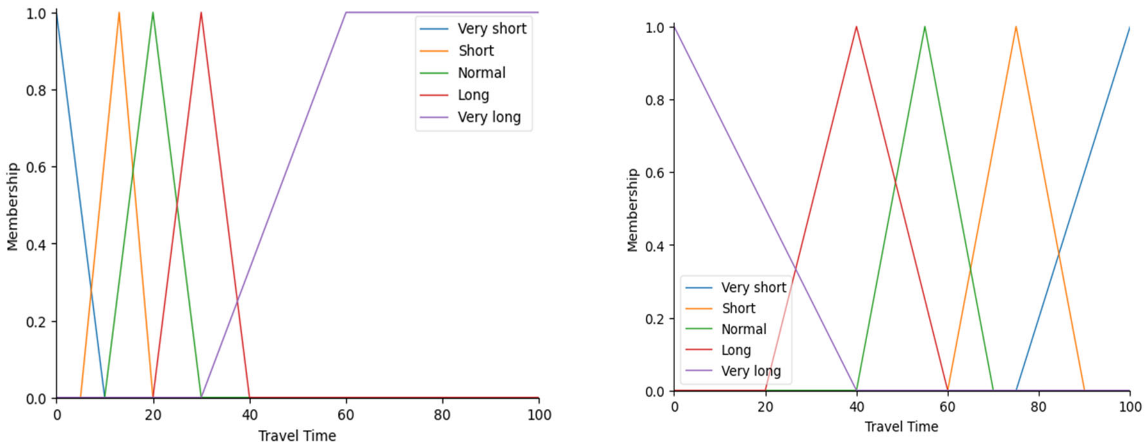

Figure 17.; *Fuzzy representation of the Travel Time parameter based on the percentage difference of the travel time of a route to the travel time of the shortest route between the same O-D pair, the difference (left) and fuzzy value according to the difference (right)*

3.5. Model Tuning

One of the most significant disadvantages of fuzzy logic is that it does not have a systematic approach and is based on personal experience and knowledge as well as tuning the output values of the fuzzy model by comparing them with real data. Although it provides flexibility for the solution of the problem, how to validate the results is a crucial question. The lack of a systematic model validation approach requires the building of model validation options suitable for the problem for the validation of fuzzy logic models.

In the route choice problem addressed within the scope of the thesis, it is assumed that the utility provided to the user by the routes chosen by the participants who participated in the "Fietstelweek 2016" event and whose GPS data were analysed is higher than alternative routes. Therefore, it is presumed that the participants choose the routes they follow. Accordingly, the attractiveness score of followed routes as a result of fuzzy modelling and inference of traffic safety, travel time, and environmental factors parameters should be greater than the alternatives of the route followed, primarily the shortest route in terms of travel time. In other words, if a route provides a greater attractiveness score to the road user than its alternatives in terms of traffic safety, environmental factors, and travel time, that route should be chosen by a road user. In order to determine which rules should be used to combine these three factors and which fuzzy set values should be considered, the value obtained as a result of fuzzy inference of the route followed is compared with the best route alternatives in terms of travel time, environmental factors, and traffic safety parameters. It can be concluded that the fuzzy logic model is valid if at least 80% of the attractiveness values of the routes followed by the participants are greater than all of the route alternatives compared.

How travel time is handled in a fuzzy logic analysis model for an O-D pair has been discussed above. On the other hand, it has not been emphasised how to use environmental factors and traffic safety for a route and generate a value since they are assessed on a link basis. In order to obtain the traffic safety and environmental factors values of a route and use them in the fuzzy logic model to be built, the traffic safety and environmental factors values of each link forming the route are multiplied by the length of each link. The sum is divided by the total length of the route to obtain two separate values for environmental factors and traffic safety that can be used in the fuzzy logic model.

$$TS = \frac{\sum_{i=1}^{n} TS_i * l_i}{\sum_{i=1}^{n} l_i}, EF = \frac{\sum_{i=1}^{n} EF_i * l_i}{\sum_{i=1}^{n} l_i} \tag{3.2}$$

In Equation 3.2, TS and EF represent the traffic safety and environmental factors parameter, respectively, of a selected route between an O-D pair. l_i represents the length of each link forming the route, while TS_i and EF_i represent the traffic safety and environmental factors parameter, respectively, of the i^{th} link.

The routes chosen by the cyclists participating in the event are analysed with a fuzzy logic model consisting of 5*3*3 = 45 rules composed of travel time, traffic safety, and environmental factors as explained above. The routes chosen in the event for the O-D pair are compared with the best alternatives in terms of traffic safety, travel time and environmental factors. Since there is no systematic approach to be followed for the fuzzy logic model, a threshold is set for the membership degrees of the parameters and the rules to be determined. The rules and the membership degrees of the parameters are tuned until the threshold is met, and the fuzzy logic model has been finalised. According to the threshold set, the utility value of at least 80% of the routes followed must have a value greater than or close to the route alternatives in the set of fastest routes. Likewise, the same threshold must be satisfied for the safest and most appealing route alternatives in the set of route alternatives.

3.6. Model Validation and Utility to Probability

In order to validate the rules and membership functions in the fuzzy logic model, the model will be applied in Hengelo, which is very similar to Enschede in terms of economic, social, network size, population, and urbanism and is located approximately 10 km west of Enschede. Thus, the validity and accuracy of the fuzzy logic model tuned for the city of Enschede will be tested in another residential area with similar characteristics. Since the fuzzy logic model is directly dependent on the defined rules, it is very difficult to generalise easily compared to other mathematical and statistical methods. Therefore, instead of choosing a very different residential area in terms of location and demographics, Hengelo, which has great similarities with Enschede, is preferred.

After the tuning and validation of the model, the probability of choosing a route alternative within the alternative set will be estimated by using the utility values obtained as a result of the model. Although a method using fuzzy logic to find the probability of choosing a route has been developed, there are no studies in the literature on the use of fuzzy logic to write probability formulae. While Teodorovic and Kikuchi (1990) conducted a study in which the probability of choosing one route over the other was calculated by using the difference in benefits between the two routes, it was concluded that this approach is not effective when the number of alternatives increases. Lotan and Koutsopolous (1993) carried out a study to estimate the probability of choosing of routes using fuzzy logic in the case of more than two alternatives, but the methodology of the study is not generalisable to this study due to the network-wide nature of this study. Therefore, in order to find the probability of choosing the routes within a set of alternative routes by using utility values, the logit formula, which is widely used and quite simple to apply, will be used. The C-logit formula, which is a variation of the logit formula with the addition of a commonality factor, will be used since it has the property of decreasing the utility of overlapping routes and increasing the utility of independent routes, and hence their choice probability. The main reason why the C-logit formulation will be specifically utilised is to reduce the probability of choosing the shortest route due to the large number of overlapping link lengths between the chosen routes and the shortest routes by utilising the C-logit method to increase the probability of choosing an independent route. In this context, maximum likelihood is used,

$$P_{ij}^{\lambda_1} = \frac{e^{\beta V_{ij}^{\lambda_1} - cf_1}}{\sum_{i=1}^{J} e^{\beta V_{ij}^{\lambda_i} - cf_i}} \tag{3.3}$$

In Equation 3.3, the coefficient β, which represents the parameter controlling dispersion in mode choice, will be found. V_{ij} represents the utility of a route and will be calculated with a fuzzy logic model. cf_k represents the commonality factor, and the utility of the overlapping routes is calculated with a fuzzy logic model and subtracted from the utility of the entire route. Thus, for any set of route alternatives between any O-D pair, the probability of choosing each alternative can be calculated.

4. RESULTS

Since fuzzy logic does not have a systematic approach and a validation method, the utility values of the routes are obtained by using rule sets defined by using travel time, traffic safety, and environmental factors parameters. In order to finalise the rule set for the fuzzy logic model, according to the criterion defined, the utility of at least 80% of the routes followed between all O-D pairs must have a utility value close to or larger than the utility values of the fastest route alternatives. The criterion is valid not only for the fastest route alternatives but also for the safest and most appealing route alternatives. That is, the utility value of at least 80% of all routes followed must have a utility value close to or larger than the utility value of the fastest, safest, and appealing route alternatives individually. Therefore, the fuzzy logic model is tuned until the specified threshold is met. Before constructing the fuzzy logic model consisting of three parameters, these three parameters are analysed by considering two parameters each. Travel time utility is calculated by considering the difference in travel time between the shortest travel time on a route basis and the route followed. Traffic safety and environmental factors are evaluated on link basis and a utility value is determined for any route with the contribution of the links in the route. Traffic safety and environmental factors parameters are evaluated with a fuzzy logic model and the attractiveness of each link forming the cycle network in terms of these two factors will be mapped. Considering the three parameters, the fuzzy logic model is finalised by trial and error, and the probabilities of choosing the route alternatives between the O-D pair of any route are found with the help of the logit formula. Thus, with a fuzzy logic model, the utility values of route alternatives for any O-D pair in a cycling network can be found and based on the utility, the probabilities of route alternatives being chosen by a cyclist are obtained.

In Enschede, a total of 302 participants took part in the event. The routes that are not considered for the fuzzy logic model are routes with a total travel time of less than 4 minutes and routes with a total travel time of more than 75 minutes. For routes under 4 minutes, although the travel time differences considered in the fuzzy logic model are small, the percentage differences are quite large due to the very short travel time. For example, if the alternative of a route with a travel time of 3 minutes is 4 minutes, although the travel time difference is 1 minute, the percentage difference is 33% and this percentage difference in the fuzzy logic model expresses the difference between the two routes as long/very long. In reality, it is assumed that a 1-minute difference would not be considered as long/very long by a road user at such short distances. For this reason, it is assumed that the percentage expression of the differences in travel time for journeys under 5 minutes is not significant, and these trips are not included in the model. For trips longer than 70 minutes, the travel times of the safest and most appealing routes differ significantly from the travel time of the shortest route for the same O-D pair. Hence, for such long trips, the attractive and safe route alternatives are not considered as realistic route alternatives because they have too long travel times to be followed by any route user. For this reason, journeys under 4 minutes and some journeys for outliers were not included in the analysis model, and a fuzzy logic model was built with the data of 235 participants in total.

Fuzzy logic values have been obtained using the Scikit-Fuzzy library. Scikit-Fuzzy is a Python-based open-source fuzzy logic library developed by the SciPy community, which produces tools for scientific and engineering-based computing. It contains many basic features of fuzzy logic and is also a useful tool for advanced applications such as fuzzy clustering (SciPy, 2023).

In the following sections, fuzzy logic models are constructed and evaluated by considering Traffic Safety & Environmental Factors, Travel Time & Traffic Safety, Travel Time & Environmental Factors and Travel Time & Traffic Safety & Environmental Factors, respectively. The purpose of the analysis models using two parameters is to make it easy to understand the effect of each parameter by comparing it with the analysis model using three parameters. It is also aimed to demonstrate how an extra parameter changes the rules that need to be defined in the analysis model and the results obtained.

4.1. Traffic Safety & Environmental Factors

Since traffic safety and environmental factors parameters contain three fuzzy sets each, 3*3=9 rules are used in the inference system to be built. These rules are blended with 'AND' conjunction and a rule set consisting of 'IF-THEN' block is defined. The value obtained as a result of the defined rule sets and the defuzzification is 'Attractiveness'. 'Attractiveness' is defined by five different fuzzy sets. Five different fuzzy sets are labelled as 'very high, high, medium, low, very low', and the fuzzy numbers showing the membership degrees of the fuzzy sets are shown in Figure 18.

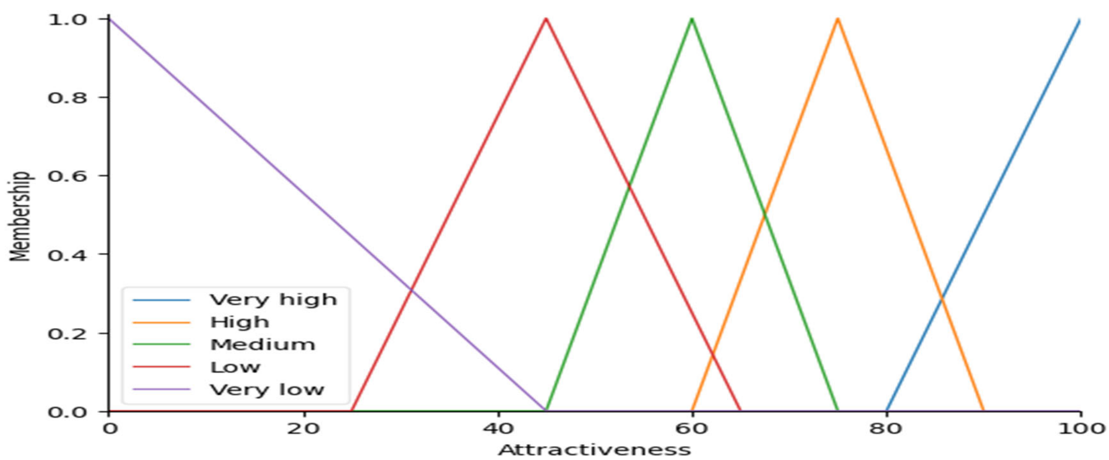

Figure 18.; *Fuzzy number membership function of Attractiveness Score, the result of fuzzy logic inference system*

According to the defined rule set, if a link is both safe and attractive, the attractiveness score is very high, or if a link with normal traffic safety has low attractiveness, the attractiveness score is considered as low. Table 4 shows the rules defined using traffic safety and environmental factors parameters and their inference.

Table 4.; *Rule set and inference system consisting of traffic safety and environmental factors*

Rule Number	IF	Traffic Safety	AND	Environmental Factors	THEN	Attractiveness (Utility)
Rule 1	IF	Safe	AND	Repellent	THEN	High
Rule 2	IF	Normal	AND	Repellent	THEN	Low
Rule 3	IF	Unsafe	AND	Repellent	THEN	Very Low

Rule 4	IF	Safe	AND	Neutral	THEN	High
Rule 5	IF	Normal	AND	Neutral	THEN	Medium
Rule 6	IF	Unsafe	AND	Neutral	THEN	Low
Rule 7	IF	Safe	AND	Attractive	THEN	Very high
Rule 8	IF	Normal	AND	Attractive	THEN	High
Rule 9	IF	Unsafe	AND	Attractive	THEN	Medium

The results obtained as a result of the rule set and the inference mechanism are then defuzzified and a classical value is obtained. As mentioned before, a classical number value is obtained by using the centroid method for the defuzzification. Figure 19 shows the attractiveness scores of the links forming the cycling network in Enschede. Since it is not possible to display the results as fuzzy in ArcGIS used in the preparation of the map, the results are classified according to the classical set logic using more ranges of values. According to Figure 19, priority should be given to increasing the attractiveness of links with a score of 60 and below. These links have received low scores either because their traffic safety is insufficient or because their environmental factors have low attractiveness, or even both factors for some links. In order to increase the attractiveness of these links, their traffic safety should be increased since it is not possible to change the land use, which is considered under the environmental factors parameter within the scope of the study. For this reason, cycle lanes should be provided to the links that do not have cycle lanes, starting with the links with a score of 40 and below. In the links where cycle paths are not separated by physical barriers, especially in the links with heavy vehicle traffic, efforts should be made to separate the cycle path with physical barriers. Thus, the traffic safety of links with low attractiveness can be increased and they can be made more attractive for road users.

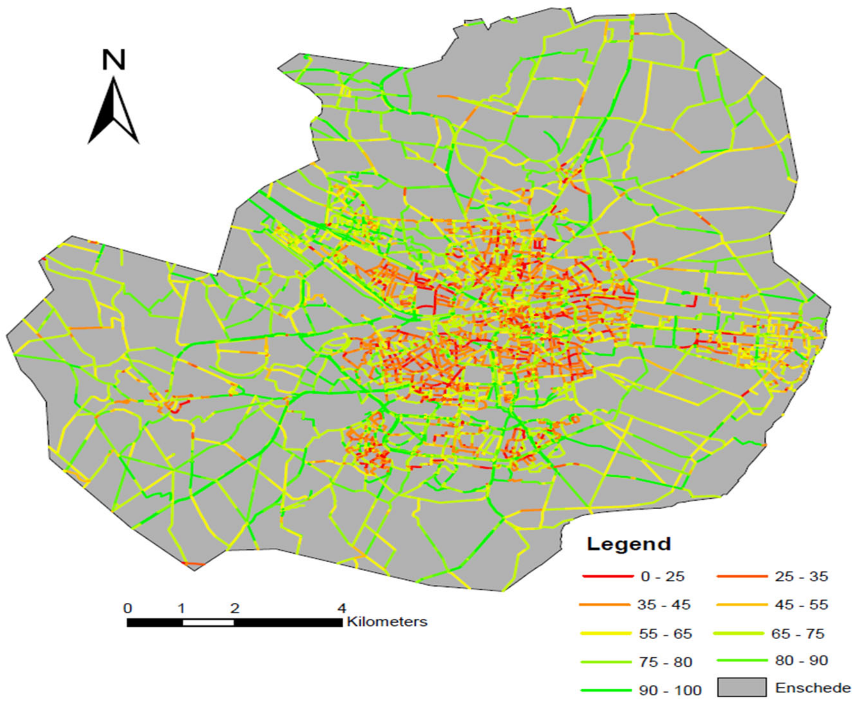

Figure 19.; *The attractiveness of the Enschede cycle network in terms of traffic safety and environmental factors*

In the fuzzy logic model built using only traffic safety and environmental factors parameters, 82.55% of the routes followed by the participants have better utility values than the shortest route. On the other hand, the utility values of the best routes in terms of traffic safety and environmental parameters are better than a significant portion of the utility values of the routes followed. Accordingly, 57.02% of the most attractive routes in terms of environmental factors have a better utility value than the followed routes, while 65.53% of the best routes for traffic safety provide a higher utility value to the road user than the followed routes. When the shortest route alternatives are compared with the safest and most appealing route alternatives, the utility values obtained are quite unfavourable to the shortest route alternatives. 30.64% of the shortest route alternatives have a better utility value than the shortest route alternatives. When the shortest route alternatives are compared with the most appealing route alternatives in terms of environmental factors, only 37.02% of the shortest route alternatives have a higher utility value.

4.2. Travel Time & Environmental Factors

Since travel time and environmental factors contain five and three fuzzy sets, 5*3=15 rules are used in the inference system to be built. These rules are blended with 'AND' conjunction as in the comparison of traffic safety and environmental factors parameters and a rule set consisting of 'IF-THEN' block is defined. The defined rule sets and the values obtained as a result of the defuzzification are based on the same 'Attractiveness' fuzzy set. Table 5 shows the defined rules.

Table 5.; *Rule set and inference system consisting of travel time and environmental factors*

Rule Number	IF	Travel Time	AND	Environmental Factors	THEN	Attractiveness (Utility)
Rule 1	IF	Very short	AND	Repellent	THEN	High
Rule 2	IF	Short	AND	Repellent	THEN	Medium
Rule 3	IF	Normal	AND	Repellent	THEN	Low
Rule 4	IF	Long	AND	Repellent	THEN	Very low
Rule 5	IF	Very long	AND	Repellent	THEN	Very low
Rule 6	IF	Very short	AND	Neutral	THEN	Very high
Rule 7	IF	Short	AND	Neutral	THEN	High
Rule 8	IF	Normal	AND	Neutral	THEN	Medium
Rule 9	IF	Long	AND	Neutral	THEN	Low
Rule 10	IF	Very long	AND	Neutral	THEN	Very low
Rule 11	IF	Very short	AND	Attractive	THEN	Very high
Rule 12	IF	Short	AND	Attractive	THEN	Very high
Rule 13	IF	Normal	AND	Attractive	THEN	High
Rule 14	IF	Long	AND	Attractive	THEN	Medium
Rule 15	IF	Very long	AND	Attractive	THEN	Low

Considering the travel time, significant differences in favour of the shortest route are observed in the results compared to the fuzzy logic model constructed with traffic safety and environmental factors parameters. Accordingly, in the fuzzy logic model constructed by considering environmental factors and travel time parameters, the utility of only 63.40% of the routes followed by the participants is larger than the attractiveness values of the shortest route alternatives. Compared to the results of the fuzzy logic model constructed with the traffic safety and environmental factors parameters discussed in the previous sub-heading, considering the travel time reduced the utility of 45 routes below the shortest route alternative. However, the followed routes have a larger utility than a significant portion of the safest and most appealing route alternatives. Accordingly, when only travelling time and environmental factors are considered for the followed routes, the utility is higher than 83.83% of the safest route alternatives and 89.36% of the most appealing route alternatives. It is observed that travelling time significantly reduces the attractiveness of the safest and most appealing route alternatives. This difference in travel times suggests that road users need to detour in order to choose the safest or most attractive routes, with a significant trade-off in travel time. The route followed by the 4 event participants and the shortest, safest, and most appealing route alternatives between the O-D pair are shown in Figure 20. The main reason for demonstrating route alternatives for a limited number of event participants is to avoid cluttering the figure due to the common use of many links in the O-D pair followed by the participants.

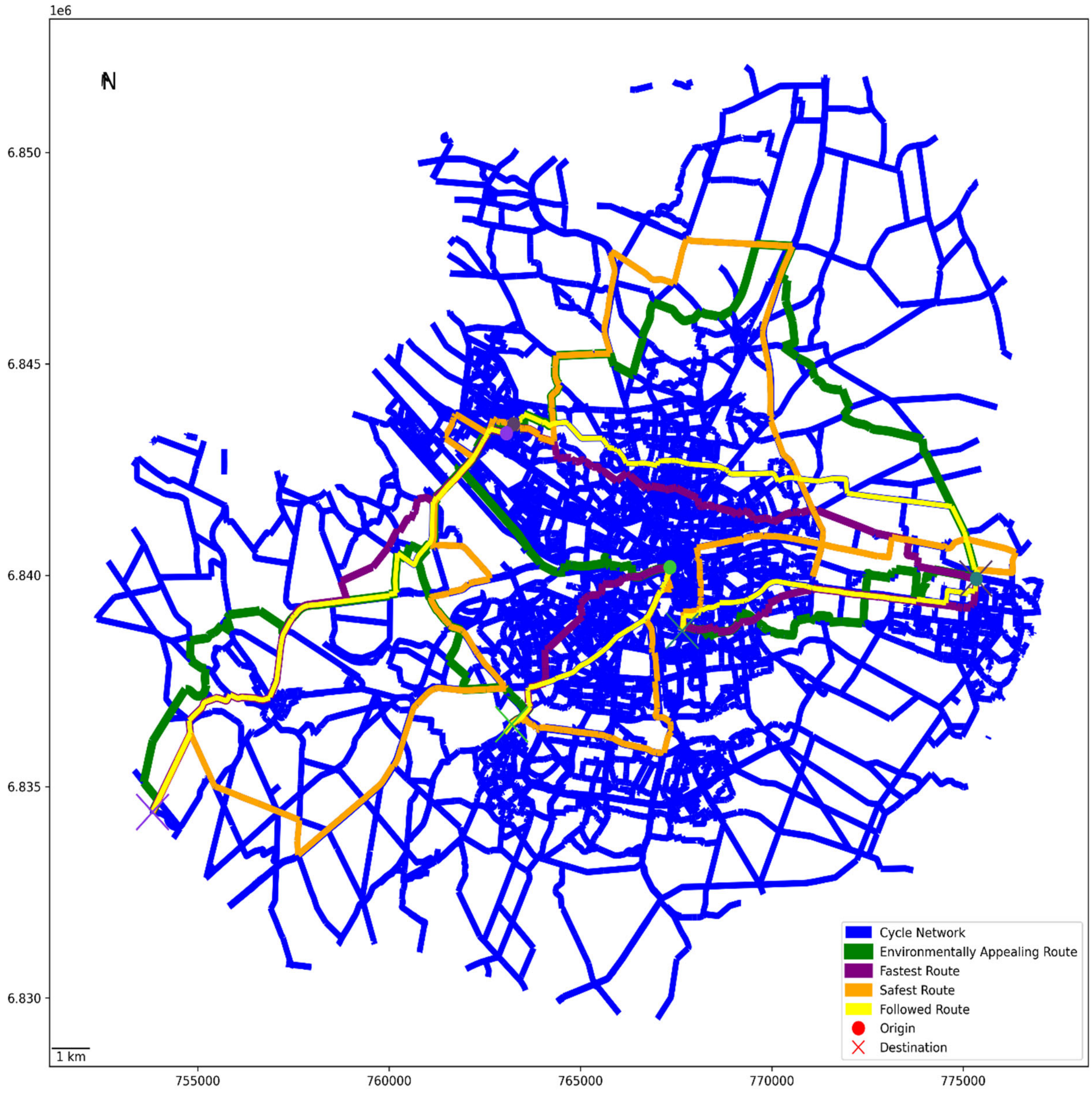

Figure 20.; *For five event participants, the route followed between O-D pairs (yellow) and the fastest (purple), most attractive (green), and safest (orange) route alternatives*

4.3. Travel Time & Traffic Safety

Since travel time is included in the analysis model with five fuzzy sets and traffic safety is included in the analysis model with three fuzzy sets, the 'Attractiveness' values of the routes followed and the best route alternatives in terms of travel time and traffic safety were found as a result of the inference system built by defining a total of 5*3 = 15 rules. Table 6 shows the rules defined in the fuzzy logic analysis model constructed with travel time and traffic safety.

Table 6.; *Rule set and inference system consisting of travel time and traffic safety*

Rule Number	IF	Travel Time	AND	Traffic Safety	THEN	Attractiveness (Utility)
Rule 1	IF	Very short	AND	Safe	THEN	Very high
Rule 2	IF	Short	AND	Safe	THEN	Very high
Rule 3	IF	Normal	AND	Safe	THEN	High
Rule 4	IF	Long	AND	Safe	THEN	Medium
Rule 5	IF	Very long	AND	Safe	THEN	Low
Rule 6	IF	Very short	AND	Normal	THEN	High
Rule 7	IF	Short	AND	Normal	THEN	High
Rule 8	IF	Normal	AND	Normal	THEN	Medium
Rule 9	IF	Long	AND	Normal	THEN	Low
Rule 10	IF	Very long	AND	Normal	THEN	Low
Rule 11	IF	Very short	AND	Unsafe	THEN	High
Rule 12	IF	Short	AND	Unsafe	THEN	Medium
Rule 13	IF	Normal	AND	Unsafe	THEN	Low
Rule 14	IF	Long	AND	Unsafe	THEN	Very low
Rule 15	IF	Very long	AND	Unsafe	THEN	Very low

As a result of the fuzzy logic model constructed using travel time and traffic safety parameters, the utility of the routes followed by the event participants is better than 73.19% of the route alternatives compared. Accordingly, when the utility of the routes followed for the road user is compared with the most appealing route alternatives, similar results are observed with the fuzzy logic model based on travel time and environmental factors. The utility of 216 of the routes followed has a value larger than the utility of the most appealing route alternatives, which constitutes 91.91% of the routes included in the analysis model. It indicates that the most appealing route alternatives are quite unfavourable in terms of travel time, and if a road user follows these route alternatives, he/she must compromise significantly in terms of travel time. Comparing the shortest route alternatives between the routes followed and the O-D pair, the utility of the followed route for 172 road user is larger. 172 routes constitute approximately 73% of the routes included in the analysis model. Finally, in 191 routes, which constitute 81.28% of all routes, the utility of the routes followed for the road user are larger than the utility of the safest road alternatives. When the results are compared with the fuzzy logic model where travel time and environmental factors are evaluated, the results show that the shortest route alternatives between O-D pairs are better than the followed routes in terms of environmental factors but provide lower utility to road users in terms of traffic safety.

4.4. Travel Time & Traffic Safety & Environmental Factors

In order to analyse the route choice with a fuzzy logic model, a rule-based fuzzy logic model has been built using three selected parameters. In the first fuzzy logic model, travelling time was represented by four fuzzy sets in order to define fewer rules. However, the use of four fuzzy sets caused various inconsistencies in terms of results. The most significant of these inconsistencies is that the comparison of the travel time with the shortest travel time is represented by two fuzzy sets as 'long' and 'very long', while the travel lengths close to the shortest travel time are represented by only one fuzzy set 'short'. Especially for routes up to 15% longer than the shortest travel time, the utility values were lower than expected due to the single set representation. Therefore, when defining the travel time, the routes described as 'long' and 'short' were represented by two fuzzy sets in order to eliminate the situation that occurred to the disadvantage of the routes described as 'short' while obtaining the utility values in the modelling. Although the increasing number of fuzzy sets in fuzzy logic models increases the number of rules that need to be defined in the rule set to be defined, it is only possible to obtain more accurate utility values as a result of the inference mechanism by constructing a sufficient number of fuzzy sets. In the fuzzy logic model built by considering three parameters, the inference mechanism has been formed by defining 5*3*3 = 45 rules. These defined rules are given in Appendix 1 and the code block showing how the defined rules are applied in Fuzzy Scikit is given in Appendix 2.

Since there is no specific systematic to be followed in the fuzzy logic model, the rules can be defined arbitrarily. In order for the rules not to be defined arbitrarily and for the defined rules to create a coherent whole, the criteria should be chosen, and the results obtained should be compared with this criterion, and the rules should be changed if the results are not consistent. In this context, it has been aimed to obtain a utility value greater than the utility values of 80% of the shortest, safest, and most attractive route alternatives of the utility values obtained for the routes followed by the participants of the event with the defuzzification of the fuzzy numbers defined within the framework of the defined rules. Nine fuzzy logic models built for this purpose were not successful. Four of the fuzzy logic models produced less accurate results as a result of representing the travel time as a result of being represented by four different fuzzy sets. The other five analysed models, represented by five different fuzzy sets, all showed a utility value larger than at least 80% of the most appealing routes from an environmental point of view. In three of the five models, the 80% threshold for comparing the utility values of the followed routes with the safest route alternatives could not be exceeded. For the shortest route alternatives, the utility values obtained as a result of the fuzzy logic models tested had a lower utility than the routes followed for at least 68% and at most 78% of the routes and remained below the 80% threshold. In the tenth fuzzy logic model tested, the utility values obtained from the routes followed had a higher utility value than at least 80% of the shortest, safest, and most appealing route alternatives. Accordingly, 81.28% of the routes followed in the finalised fuzzy logic model are larger than the shortest route alternatives between O-D pairs for the road user. When the utility values of the routes followed are compared with the safest route alternatives, it is observed that 83.83% of the routes followed have larger utility values than the safest route alternatives. On the other hand, only 12.34% of the most appealing route alternatives from an environmental point of view had larger benefits for the road user than the followed routes. For 206 routes, the routes followed provided higher utility to the road user compared the most appealing route alternatives. Figure 21 shows a total of 171 routes where the highest benefit belongs to the followed route as a result of the comparison of the benefit values of the followed route, the shortest route, the safest route, and the most appealing route alternatives. The origin and destination of each route are marked with a different colour.

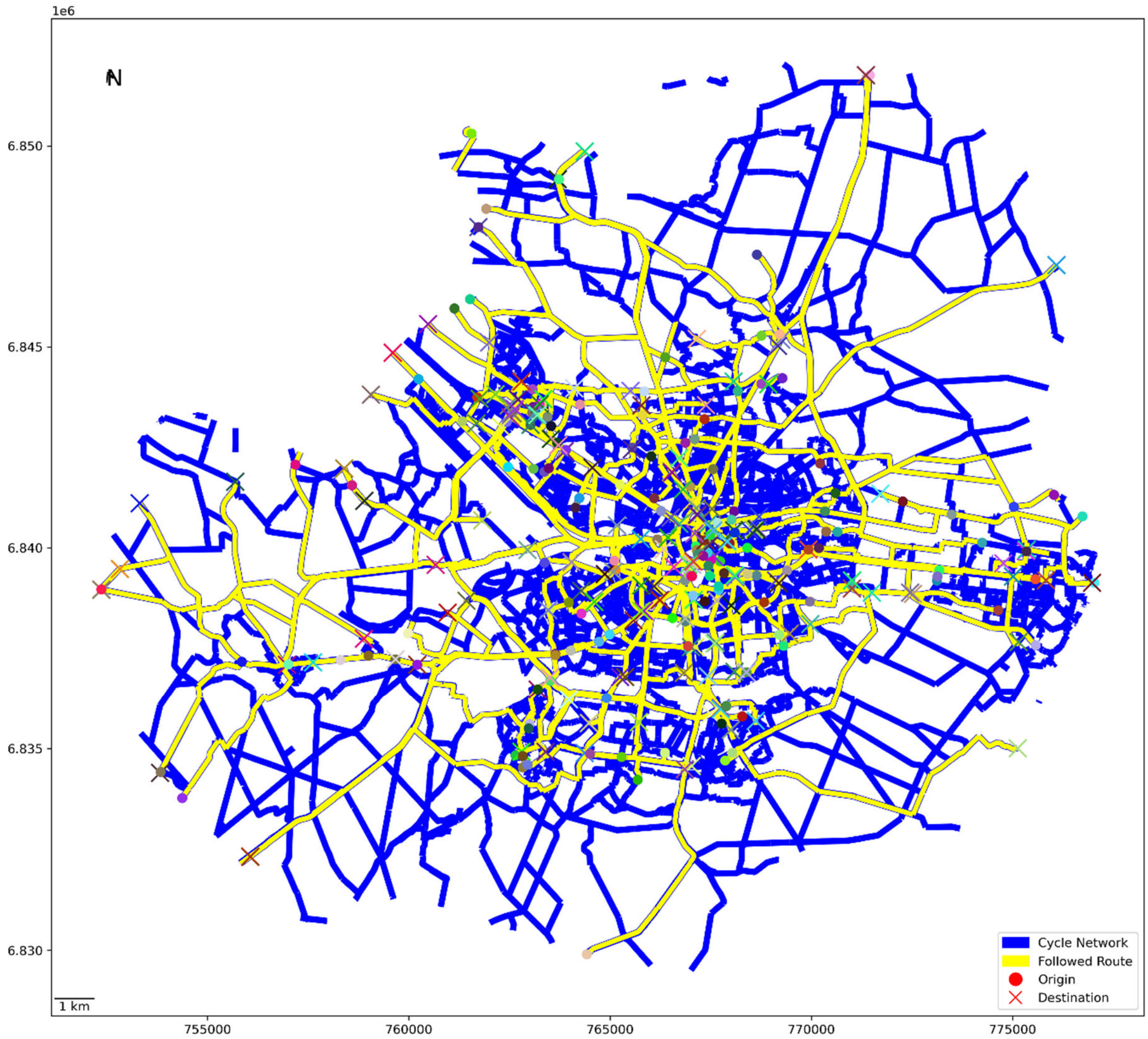

Figure 21.; *Routes where the route followed provides more utility to the road user than the shortest, safest, and most appealing route alternatives*

When the utility of the four route alternatives for the road user is compared according to the fuzzy logic model, the routes chosen by the event participants are the routes with the highest utility value among the alternatives in 73% of all routes. For 6 routes, which constitute approximately 2.50% of all routes, the most appealing route provides the highest utility to the road user among the route alternatives. Similarly, for 13 routes, which constitute approximately 5.50% of all routes, the safest route is the route with the highest utility value. For 44 routes, corresponding to 18.72% of all routes, the shortest route alternative has a larger utility value for a road user compared to the other route alternatives. In Figure 22, the routes with the highest utility value among the alternatives are shown in three different colours. The routes shown in orange are the routes where the safest route alternatives have a greater utility value than the other route alternatives. Safe route alternatives generally have greater utility value on longer routes in terms of travel time. The routes shown in green are the routes where the most environmentally attractive routes provide the largest utility value. For routes with a travel time between 5-8 minutes, the utility values of the most appealing route alternatives are larger compared to their alternatives. The routes shown in purple are the route alternatives where the shortest route alternatives provide the highest utility.

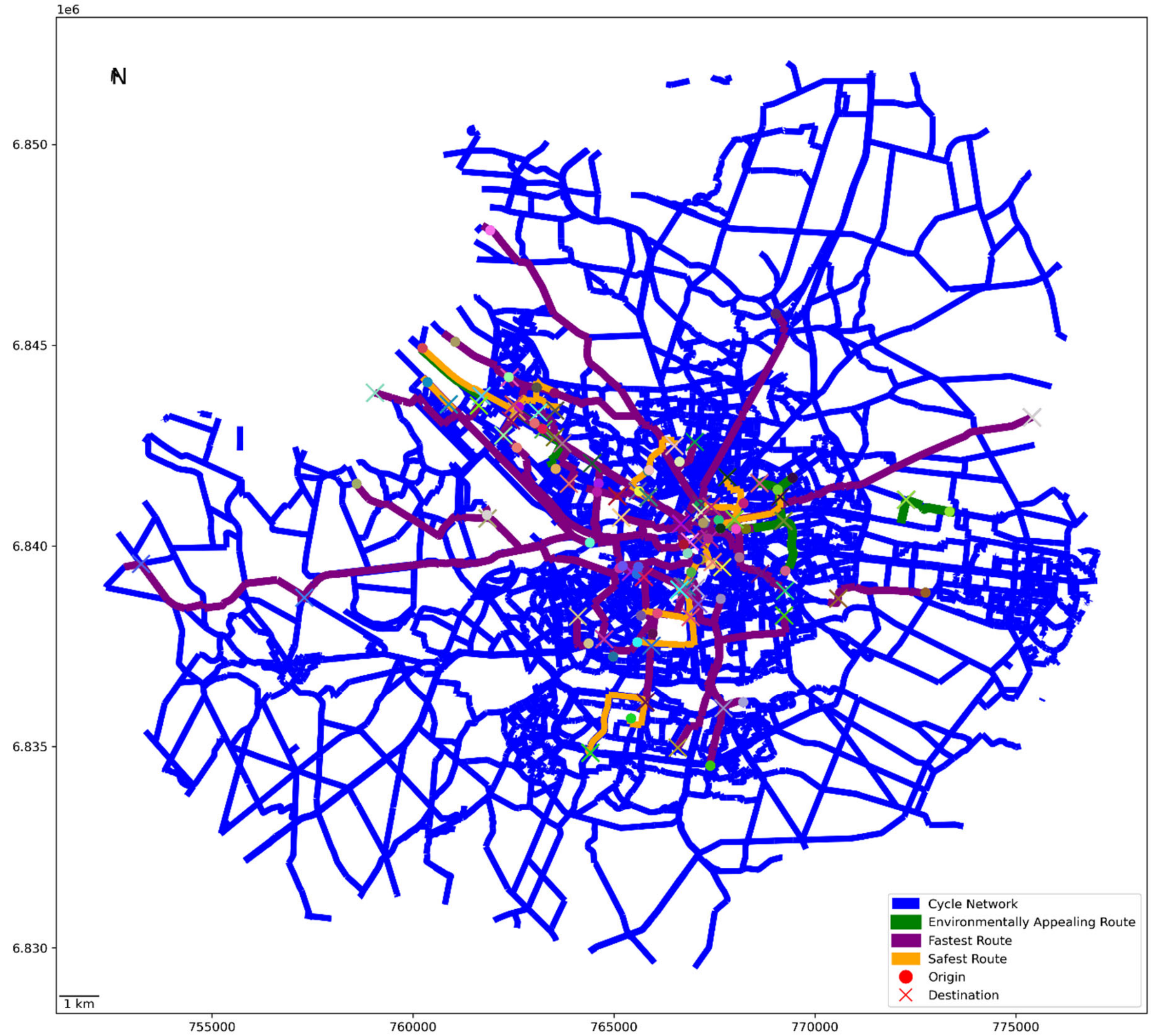

Figure 22.; *Route alternatives with the highest utility values among the O-D pairs travelled by the event participants*

4.5. Results in Hengelo

Although there are 215 cyclists participating in the event in Hengelo, the fuzzy logic model is analysed using the data of only 136 participants. 57 of the participants in Hengelo are not included in the analysis model because their travel time within the scope of the event is less than 5 minutes. The remaining 22 routes are not included in the analysis model because the route followed completely overlapped with one of the fastest/safest/most attractive routes, and the time differences between the route alternatives are significantly larger. According to the

136-participant fuzzy logic analysis model, 81.62% of the routes followed have a greater utility for the road user than the shortest route alternatives. When the utility values of the safest and most appealing route alternatives are compared with the utility values of the routes followed, this is larger than both the proportion of the shortest route alternatives and the proportion of the same route alternatives in Enschede. The 122 routes followed by event participants in Hengelo, which corresponds to approximately 89.71%, have utility values larger than the safest route alternatives. For a total of 117 routes, corresponding to 86.03% of all routes, the utility values of the routes followed are larger than the utility of the most appealing route alternatives. Figure 23 shows the attractiveness of the links forming the Hengelo network in terms of traffic safety and environmental parameters. Since it is not possible to display the results as fuzzy in ArcGIS used in the preparation of the map, the results are classified according to the classical set logic using more ranges of values.

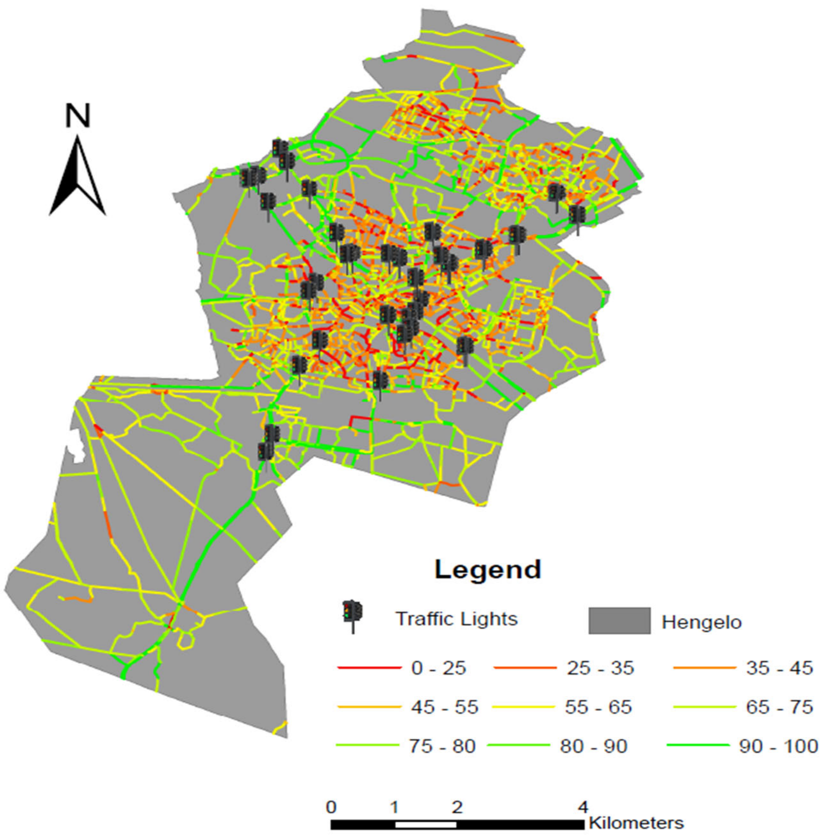

Figure 23.; *The attractiveness of the Hengelo cycle network in terms of traffic safety and environmental factors*

For 73.53% of the set of route alternatives consisting of four route alternatives, which means 100 routes, the utility values of the routes followed are the largest. For 31 routes, which constitute 22.79% of all routes, the shortest route between O-D pair has the largest utility value. For a total of 5 routes, the safest or most attractive route alternative between O-D pair has the largest utility value. Figure 24 shows the route alternative with the largest utility value between O-D pairs.

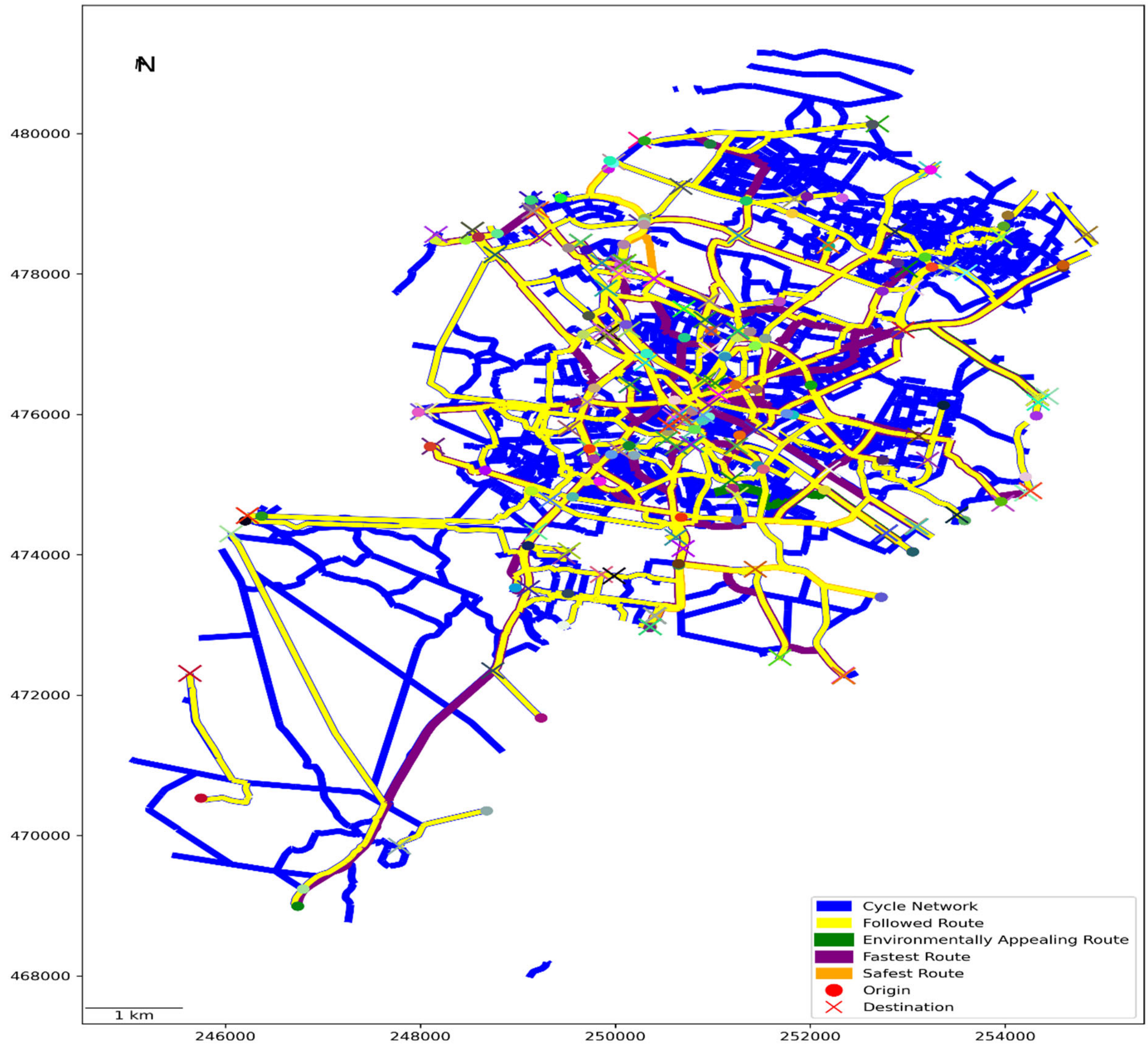

Figure 24.; *Routes with the highest utility values among the O-D pairs travelled by the event participants in Hengelo*

According to the fuzzy logic model, a route chosen by a road user should be larger than its alternatives in terms of the utility value obtained from the combination of travel time, traffic safety, and environmental factors parameters. The greater the algebraic difference between the utility value obtained for a route and the utility values of the routes in the alternative set, the higher the probability of choosing that route compared to its alternatives. According to the results of the fuzzy logic model established in this context, there are 172 routes in Enschede, which corresponds to 73.19%, and 100 routes in Hengelo, which corresponds to 73.53%, with the highest probability of being chosen. It demonstrates that the fuzzy logic model is able to correctly predict about 73% of the route choices in the analysed "fietstelweek" dataset. The approach to probability calculations is discussed in the following heading.

4.6. Utility to Probability

By using the utility functions found, the probabilities of choosing the routes in the alternative set are estimated using C-logit. The point where the C-logit model differs from the classical logit model is the common links in the routes in the alternative set. The total utility of the common links of the routes in the set of alternatives to the user is calculated fuzzily. Then, the commonality factor, cf_k, is calculated using Equation 4.1.

$$cf_k = \beta_0 ln \sum_{h \in I_{rs}} \left(\frac{V_{hk}L_{hk}}{(V_h L_h)^{0.5}(V_k L_k)^{0.5}}\right)^{\gamma} \tag{4.1}$$

In Equation 4.1, V_{hk} is the total utility of the common links of routes h and k, L_{hk} is the total length of the common links of routes h and k, L_h and L_k are the total lengths of routes h and k, respectively, and β_0 and γ are positive parameters. After calculating the utility of the common routes, the probability of choosing each route within a set of route alternatives is calculated using Equation 4.2.

$$P_{ij}^{\lambda_1} = \frac{e^{\beta V_{ij}^{\lambda_1} - cf_1}}{\sum_{i=1}^{J} e^{\beta V_{ij}^{\lambda_i} - cf_i}} \tag{4.2}$$

However, since the β value is not known at this point, the β value is calculated using maximum likelihood. Since the route alternatives for each O-D pair and the route selected among these alternatives are known, the β value that maximises the probability for 235 routes is calculated and is found 1/19. Thus, the probabilities of choosing the fastest, safest, and most appealing route alternatives for each O-D pair are calculated by excluding the route followed from the set of alternative routes. Table 7 shows the utility values and choosing probabilities of the route alternatives for the five O-D pairs in Enschede.

Table 7.; *Utility values of the alternative routes between the five O-D pairs in Enschede, and their probabilities of being chosen according to the C-logit and logit formula*

Participant No	Shortest Route Utility	Safest Route Utility	Appealing Route Utility	Pr C-logit (Shortest Route)	Pr C-logit (Safest Route)	Pr C-logit (Appealing Route)
1	78.72	76.84	34.74	47.10%	46.65%	6.25%
2	91.00	91.00	83.20	35.55%	34.95%	29.51%
3	79.77	59.17	62.33	55.66%	21.92%	22.42%
4	78.18	37.24	45.00	76.00%	9.33%	14.67%
5	91.94	84.90	78.29	46.70%	39.20%	14.10%
Participant No	Shortest Route Utility	Safest Route Utility	Appealing Route Utility	Pr Logit (Shortest Route)	Pr Logit (Safest Route)	Pr Logit (Appealing Route)
1	78.72	76.84	34.74	48.89%	45.18%	4.93%
2	91.00	91.00	83.20	37.55%	37.55%	24.90%
3	79.77	59.17	62.33	57.55%	19.46%	22.99%
4	78.18	37.24	45.00	77.50%	8.98%	13.52%
5	91.94	84.90	78.29	45.92%	31.70%	22.38%

5. DISCUSSION

A study has been carried out on the use of fuzzy logic model on a network and its predictive power in order to predict the utility of alternatives, which is frequently used in prediction models to estimate which route a road user may prefer. Unlike similar studies aiming to estimate the utility of a route with fuzzy logic, GPS data is used instead of a survey. Due to the structure of the data used, it is aimed to build a fuzzy logic model that can be valid for any O-D pair in the whole network. In the study, a constraint is set in order to determine the membership functions and fuzzy logic rules that should be defined in the fuzzy logic model, since no information about the route choice behaviour of road users can be obtained through a survey and direct real-world data is used. With the constraint, the rules and membership functions of the fuzzy logic model are tested, and the model is finalised.

According to the fuzzy logic model, travel time is the most influential parameter among the three parameters for maximising the utility of cyclists, while the least influential parameter is the environmental factors parameter. The utilities of the routes followed have a utility value larger than 81.28% of the fastest route alternatives, while the same proportions are 83.83% for the safest road alternatives and 87.66% for the most appealing road alternatives, respectively. When the utilities of the routes in a four-alternative option set consisting of the route followed and the best route alternatives in terms of three parameters are compared, the route followed for 171 routes, the fastest route alternative for 44 routes, the safest route alternative for 14 routes, and the most appealing route alternatives for 6 routes have the largest utility. From this point of view, it can be said that only 73% of the routes followed in the fuzzy logic model have the largest utility value among the route alternatives. Compared to the results of three studies in the literature using fuzzy logic models, the result of the study has a low predictive power. The fuzzy logic models developed by Murat and Uludağ (2008), Dubey et al. (2013) and Dhulipala et al. (2020) in survey-based studies for only one O-D pair were able to predict the probability of a road user's choice of road alternatives in a route set consisting of three or four route alternatives between an O-D pair with an accuracy of 90% or more. The main reasons for the high difference in prediction power are the set of alternative routes, the number of rules used, the methodological difference in the study, and the application of the results to the whole network instead of a specific O-D pair. These reasons for the difference in practice and suggestions for future studies to minimise the effects of these reasons will be discussed below.

The set of route alternatives used in this study consists of the safest, fastest, and most appealing routes between the O-D pair followed by the road user according to GPS data. As such, it is questionable to what extent the route alternatives for any O-D pair represent actual or potential route alternatives. Even the shortest route in terms of travel time between an O-D pair may not always be a direct route alternative for various reasons, such as road safety, traffic congestion, necessity of detouring from the route alternative for a while (Duckham & Kulik, 2003). Since the shortest route alternative always scores the highest in terms of travel time by definiton, it is not possible to obtain a utility value higher than the utility of the shortest route alternative unless it scores below the average in terms of safety and environmental factors parameters, or the travel time of the route alternative followed is close to the route with the shortest route alternative. In addition, for the most appealing route alternatives for O-D pairs with a travel time longer than 25 minutes and for the safest route alternatives for a travel time longer than 40 minutes, the travel time deviates significantly from both the shortest route alternative and the travel time of the route followed. This situation leads to a low fuzzy value of the route alternatives in terms of travel time, while the total utility value obtained as a result of the fuzzy logic model decreases. Widely used navigation applications, such as Google Maps, Waze, Apple Maps, etc. list route alternatives with close travel times to a road user while listing route alternatives. Therefore, a set of route alternatives consisting of routes with close travel times may be more effective in terms of both the realistic alternatives of the route alternatives and the accuracy of the utility values generated by the fuzzy logic model. When evaluating the utility of route alternatives in terms of travel time, a similar approach to the one used in this study can be adopted. Instead of taking the travel time of the route with the shortest travel time as a basis, the utility of other route alternatives in terms of travel time can be evaluated by taking the travel time of the route alternative with the shortest travel time in the set of route alternatives as a basis. In short, instead of the set of route alternatives consisting of the best route alternatives in terms of the parameters between an O-D pair, the use of route alternatives given by navigation applications actively used by many road users in transportation will contribute both to increase the explanatory power of the obtained results and to obtain the utility values of these routes by evaluating realistic route alternatives.

Another factor affecting the explanatory power of the fuzzy logic model is the number of fuzzy sets used to represent the parameters included in the model. In general, the more fuzzy sets used to describe a parameter in a model, the more accurately the complexity of the relationship between inputs and outputs can be captured (Trillas & Eciolaza, 2015). The use of more fuzzy sets allows the model to represent a wider range of inputs, as well as to better account for uncertainty and imprecision in the relationship between inputs and outputs. The travel time used in the fuzzy logic model is represented by five fuzzy sets, traffic safety and environmental factors parameters are represented by three fuzzy sets each, and the output of the fuzzy logic model, attractiveness (utility), is represented by five different fuzzy sets. The representation of traffic safety and environmental factors parameters with three fuzzy sets may have caused the route alternatives to have less or more utility values than they were supposed to be as a result of the fuzzy logic model. The representation with three fuzzy sets affected the values of the membership function used to separate safe route-normal-unsafe route for traffic safety. Considering the mathematical application of the centroid method, which is used as a fuzzy logic model, small changes in rules or membership functions can cause relatively large effects in the results when the fuzzy sets are small. For this reason, representing the traffic safety and environmental factors parameters with five fuzzy sets instead of three will both increase the precision of the utility values of the routes and contribute positively to the accuracy of the model's explanation of the results. However, this increase in the number of fuzzy sets will require the definition of 125 rules instead of 45 rules. This will complicate the model and increase the time to be spent for changes in the rules in case the desired result is not obtained. At this point, Adaptive Neuro-Fuzzy Inference System (ANFIS) can be used as an alternative method. ANFIS is a model based on a hybrid learning algorithm that automatically adjusts the membership functions and rules of the fuzzy logic inference system based on a set of input-output training data. With ANFIS, the rule system to be defined to obtain the utility value and the membership functions of the parameters included in the model can be derived automatically. However, ANFIS may require large training data in order to learn the parameters of the fuzzy inference system and neural network accurately. It contradicts the aim of developing an alternative model based on less data. Nevertheless, ANFIS can be considered as an alternative to fuzzy logic for fuzzy models that will be built using a small number of parameters as in this study. Thus, the time spent in building and tuning the fuzzy logic model can be significantly reduced.

Another important reason for the low performance of the methodology followed in the study in explaining the results compared to similar studies is the methodological difference. In similar studies, a fuzzy logic model was built by using the expectations and opinions of road users by conducting a survey, while in this study, a fuzzy logic model was built by reverse reasoning about the reasons why users may have chosen the route they followed. Fuzzy logic is a method used for modelling problems where it is difficult to make an absolute separation between categories, based on subjectivity and containing various uncertainties. The success of the fuzzy logic model depends on the degree of membership, which is defined as the degree of membership and is defined as the conformity of a variable to a fuzzy set, and the rules defined between the fuzzy sets to which the variable is connected with a degree of membership. There is no specific systematic in obtaining the membership degrees and defining the rules. In addition, the definition of the utility of a route varies for each road user. For some road users, the safety of the route may come to the forefront, while for others the shortest possible travel time may be reason enough to put safety concerns aside. A small survey consisting of 8-10 questions could have been conducted with the event participants to better understand their reasons for choosing the routes they followed and the factors affecting their cycling and route choice behaviour. Thus, the utility of a cyclist's chosen route could be found more accurately by taking into account various situations, such as whether a parameter other than the three factors suggested based on the literature review should be used, membership functions, the rules by which fuzzy sets should be combined, how the parameter used should be weighted, and the weighting of the participant's travel purpose in the model. Furthermore, utilising the views and opinions of road users would have resulted in less trial and error in the finalisation of a fuzzy logic model, resulting in less time loss and a model with more explanatory power. In this context, instead of utilising only real-world data, it would be useful to take the opinions of road users for the building and results of the fuzzy logic model developed with a subjective approach.

One of the biggest advantages of the fuzzy logic model proposed in the study is the utilisation of open-source data without an active data collection process. Other advantages of the model can be listed as the fact that the parameters considered can be easily modelled thanks to the use of linguistic expressions, as well as the fact that it can be applied free of charge with the Scikit-Fuzzy library in the Python coding language, which is widely used. By means of the proposed fuzzy logic model, the best route alternatives can be quickly generated according to the desired parameters, and the chosen route alternatives between any desired O-D pair can be quickly evaluated. The disadvantage of the model is that the membership functions of the fuzzy sets and the rules are completely dependent on the practitioner of the model due to the lack of a systematic approach in both the construction and verification of the model, which are the general disadvantages of fuzzy logic. In order to finalise the rules and membership functions in this proposed model, the routes followed must have a utility better than at least 80% of the best route alternatives for the selected parameters. By using different approaches, the model can be finalised by setting various constraints for the fuzzy logic model according to the data type and amount of data.

The proposed model aims to propose an alternative model for small and medium-scale networks in developing countries with limited data that is devoid of mathematical complexity and can be understood with basic engineering education. Although the cycle network considered in this study does not constitute a direct example for developing countries, various approaches used in this study can also be used for vehicle networks. For the travel time, which is one of the most crucial parameters in route choice, a fuzzy set can be constructed as a percentage difference from the shortest travel time among the alternative routes between the selected O-D pair. For the traffic safety parameter, various factors, such as road pavement, traffic density, number of lanes, lane width, whether the travelled link is divided or not can be evaluated under a single parameter. For environmental factors and other parameters that are likely to be used, various other factors can be included in the model based on the needs of the project or the expectations of the project owner. OSM includes many such data in its maps. However, it should be noted that these maps have limited data reliability.

A similar study can be developed for route cost analysis in maritime transport. As a result of the melting of glaciers in the polar regions, which is one of the most visible consequences of global warming, it is becoming increasingly popular to investigate the suitability of using Arctic shipping routes as an alternative to existing sea routes (Theocharis et al., 2018). It is stated that the effects of global warming will be felt more severely in the coming decades (Dreyfus et al., 2022). Therefore, the seasonal use of Arctic routes is likely to increase, and they can be an alternative to existing shipping routes as they reduce the travelling time. However, Arctic routes have various disadvantages, such as being risky for navigation safety, inadequate search and rescue facilities, lack of ports along the polar route, and the need for experienced crews. The effects of cost changes in maritime transport (Wilmsmeier & Martinez-Zarzoso, 2009), where costs vary considerably due to seasonal conditions and fluctuations in the world economy, can be represented using fuzzy numbers and a fuzzy logic analysis model can be developed for the economic feasibility of Arctic routes.

In the fuzzy logic model proposed, it has been found that the routes followed in the fuzzy logic model have a better utility value than only 70% of the other alternatives in the route choice set, and various suggestions have been made to improve the model and make it more explanatory. There is a date mismatch between the data used in the model developed within the scope of this study. While the GPS data collected belongs to 2016, the current OSM data as of 2023 is used for the cycling infrastructure. Various improvements may have been made in the cycling infrastructure in the intervening seven years. This may be the reason for the ambiguities in some of the results obtained when looking at some route choices with the current cycling infrastructure. For this reason, paying attention to data timeliness in future projects may contribute positively to the explanatory power of the model. In addition, as mentioned above, knowing the travel purposes of road users, asking subjective questions to road users such as what they pay attention to in route choice, forming a set of route alternatives to be considered with the help of navigation applications, and finding the benefit of a route more precisely by using more rules can be listed as methods that can be used both to spend less time to tune the fuzzy logic model and to increase the explanatory power of the model.

6. CONCLUSION

Transport systems have an undeniable importance in planning the future of countries or cities, both because of their social structure serving the society and because they are one of the main locomotives of a country's development. Today, the number of vehicles in cities is increasing rapidly with the rapidly growing world economy and technological developments in industrial production techniques compared to the past decades. The increasing number of vehicles brings along economic and social problems such as congestion, air and noise pollution. In order to prevent or minimise the problems that may arise from the increase in the number of vehicles, transport infrastructure investments should be well planned. In small and medium-sized cities of developing countries, planning may not be done properly due to insufficient human resources and project budgets. Another obstacle to effective planning is the lack of infrastructure to collect the data required to be used in the planning phase in a qualified manner.

Understanding the route choices of road users in urban transport network planning is important for the optimum use of the capacity of the existing transport network, understanding the shortcomings of the existing transport network, and identifying the direction of the investments to be made. Logit and probit models are widely used in modelling the route choices of road users. Although logit and probit models have various advantages, they have significant disadvantages in the application phase. Depending on the increase in the number of parameters considered in the logit model, the number of data should also increase in order to establish a reliable model. Probit model, on the other hand, requires mathematically complex operations, and its application requires experience and qualification. For this purpose, a new model that has low data dependency, is not complex, can be used at network scale, and can be applied easily and quickly is proposed by using real world data especially for the planning of small and medium scale transport networks of developing countries. The proposed model is a fuzzy logic model that is frequently used in many other fields due to the ease of mathematical expression of linguistic expressions and ease of operation. With the proposed fuzzy logic model, it is aimed to answer the following questions within the scope of the thesis. The eight questions addressed in the research questions section are answered below, respectively.

***Q.1.* "*How can a fuzzy logic model be built and validated to model route choice and explain route choices of cyclists based on GPS data in a cycle network based on travel time, traffic safety, and environmental factors?*"**

While building the fuzzy logic model, the parameters to be used in the model are determined by the literature research and categorised under three groups. The parameters of travel time, traffic safety, and environmental factors are used in the fuzzy logic model. An inference mechanism is built by using the 'and' conjunction and the 'if-then' structure. The fuzzy number obtained as a result of the inference mechanism has been defuzzified. The number obtained as a result of the defuzzification is the utility of the route for a road user. In an inference mechanism built in a fuzzy logic model, rules can be defined arbitrarily. In order to prevent this arbitrariness, the rules have been changed until the preferred routes between the O-D pairs travelled by the road users, whose travel data were collected within the scope of the 'fietselweek' event in 2016 and used as a dataset, have a utility value higher than at least 80% of the safest, fastest, and most environmentally attractive road alternatives between these O-D pairs. After a total of 10 different rule set attempts, a fuzzy logic model has been built. Using the obtained utility values, the probability of choosing the route alternatives between any O-D pair is investigated. By substituting in the logit formula, it can be found which route is likely to be chosen among the alternatives identified between an O-D pair and with which probability. Thus, with a simple approach, a fuzzy logic model has been built using the available data set, with the fuzzy logic model constructed, the benefit of any desired route for the user has been found, and the probability of choosing the routes has been determined by comparing the utility of a route with the route alternatives.

As a result, fuzzy logic as a data analysis method stands out as an easy and fast method to apply even for those who have no experience with the method. Categorical data can be quantified, and "AND" or "OR" connectives can be easily processed without the need for a strict distinction for categorisations within a data set and without needing a large data set. Thus, a numerical value that can be used in practice is obtained for traffic safety and environmental parameters that are not easy to quantify. In addition to these advantages, fuzzy logic has significant disadvantages. The most prominent of these is that there is no systematic approach that can be followed both to build a fuzzy logic model and to validate it. This may pose a problem for inexperienced implementers when establishing an analysis model. Therefore, although the fuzzy logic model is quite easy to apply, the fact that it does not contain a systematic approach contradicts the argument of developing a method that is not complex and requires less experience, especially for inexperienced implementers.

The fuzzy logic model built within the scope of the study can be said that the utility values of only 73% of the routes followed in the "fietselweek" dataset have the highest utility value among the route alternatives. In order to increase this rate and to develop a fuzzy logic model with higher predictive power, the following suggestions can be adopted. When tuning the rule set, an analysis can be made by using the route alternatives suggested by navigation programmes as route alternatives between each O-D pair instead of the used route alternatives. Secondly, by using five fuzzy sets instead of three for traffic safety and environmental factors parameters, both the rules to be defined and the utility value to be obtained as a result of the rules can be obtained more accurately. Finally, in order to make sense of the route preferences of road users, a questionnaire consisting of questions about their route choice habits as well as the data of the routes can be made. Thus, rule sets can be built faster according to the route choice habits of road users. By using Adaptive Neuro Fuzzy Inference System (ANFIS) method, both the fuzzy sets and the rule set can be formed with less effort by considering the answers given by the road users and the routes followed. Thus, the construction of the fuzzy logic model can be accelerated, and more accurate results can be obtained by eliminating the necessity to define more rules as a result of using more fuzzy sets for a parameter.

Q.2. "Which sub-factors related to traffic safety and environmental factors affect the route choice decisions of cyclists and how?"

The factors affecting route choice for cyclists can be grouped under three categories: travel time, traffic safety, and environmental factors. Travel time of cyclists is affected by the number of traffic lights on the route, cycle path, traffic volume, and number of intersections. All factors other than cycle lanes have a negative effect on travel time. Traffic safety is directly affected by lighting, cycle lane separated from vehicle traffic, traffic safety, traffic volume, cycle lane markings, and paved infrastructure. Among these factors, lighting, separated cycle lane, traffic signalling, and marked cycle lane positively affect the safety of cyclists, while traffic volume negatively affects traffic safety. Finally, environmental factors are directly related to gradient, population density, and land use. Gradient and population density negatively affect cyclists' preferences, while different land uses have different effects. Greenery area, blue space, commercial area, and industrial area have a positive effect on cyclists' decisions, while residential area has a negative effect.

Q.3. "How are the cycle network links in Enschede characterised in terms of traffic safety and environmental factors?"

Figure 19 shows the attractiveness (utility) of the links as a result of the fuzzy logic model built using only traffic safety and environmental factors. According to the figure, while the attractiveness is low for the links surrounding the city centre, the attractiveness increases as you move towards the city centre due to the increase in traffic safety. Nevertheless, the attractiveness of the city centre is not very high. As moving towards the outer part of the city centre, the attractiveness of the links increases due to the increase in the green area and the presence of cycle paths.

Q.4. "How are the links that make up the cycle network in Enschede characterised in terms of travel time?"

According to OSM data, there are 142 traffic lights in the city. The traffic lights are marked on the links where they are located and an assumption is made about their effect on the travel time. This assumption is related to the waiting time at traffic lights. The waiting time at traffic lights in Enschede varies dynamically according to the traffic density. Therefore, it is assumed that there are different waiting times during rush hour and off rush hour. For trips between 07:00-09:00 and 16:00-18:00, when traffic is heavy, a total of 48 seconds is assumed to be lost for slowing down, waiting at a traffic light, and accelerating again, and a total of 23 seconds is assumed to be lost for the other time intervals when traffic is less heavy. On links that are not separated from vehicular traffic by a physical barrier during rush hour, cyclists are assumed to reduce their speed by 10-15% for the reason that they cannot overtake safely, and in the city centre where pedestrian density is high. In addition, cyclists are assumed to reduce their speed by 20% after 13:00 on weekends and between 17:00-20:00 on weekdays for their own and pedestrians' safety owing to crowdedness.

Q.5. "With what accuracy do the results obtained explain the GPS data obtained within the scope of 'fietstelweek'?"

The routes travelled by the volunteers participating in the 'fietstelweek' event are compared with various route alternatives between the O-D pairs travelled. The route alternatives compared are the shortest, safest, and most environmentally appealing route alternatives between each O-D pair. Comparing the routes taken with these route alternatives, 81.28% of the routes taken have larger utility for a road user than the shortest route alternatives, 87.66% of the routes taken have larger utility for a road user than the most environmentally appealing route alternatives, and 83.83% of the routes taken have larger utility for a road user than the safest route alternatives. In case the set of route alternatives for a road user participating in the event has these four route alternatives for each O-D pair, the route followed for a total of 171 road users, the most environmentally appealing route for 6 road users, the safest route for 14 road users, and the fastest route for 44 road users, the utility obtained by the fuzzy logic model is the largest.

Q.6. "If the fuzzy logic model does not explain the results with sufficient accuracy, what is the main reason? Are the rule sets or membership functions in the fuzzy logic model faulty, or are the parameters/sub-parameters considered insufficient to explain the results?"

The use of more fuzzy sets in the fuzzy logic model increases the accuracy of the results obtained. The fuzzy logic model uses five fuzzy sets to represent travel time, three fuzzy sets each to represent parameters of traffic safety and environmental factors, and five different fuzzy sets to represent the output of the model, which is the attractiveness (utility). The use of three fuzzy sets to represent parameters of traffic safety and environmental factors in the fuzzy logic model may have resulted in the route alternatives having utility values that are either higher or lower than their actual values. Since the fuzzy logic model employs the centroid method, even slight modifications in the rules or membership functions can lead to significant differences in the results, particularly when the numbers of considering parameters are small. Therefore, increasing the number of fuzzy sets from three to five for representing traffic safety and environmental factors parameters would improve the accuracy and precision of the utility values of the routes and enhance the model's ability to explain the outcomes.

Q.7. "Can a fuzzy logic model be easily generalized? If no, what are the impediments to generalizability, and how can these impediments be overcome?"

It is not possible to easily generalise the fuzzy model developed to understand the route choices of cyclists because the weights of the parameters that constitute the utility of road users are subjective. However, the influential factor here is not the generalisability of the rule set that forms the fuzzy logic model, but rather how the utility value obtained as a result of the defuzzification of the inference mechanism after the rule set is built is subjected to comparison. Although it is not possible to generalise the rule set established for the model as it is, the rule of by obtaining a final rule set by changing the best route alternatives in the sub-parameters of the rule set until they exceed the 80% threshold value can be generalised. Even though there is not a systematic approach for fuzzy logic models, determining a threshold value and defining a rule set accordingly prevents arbitrary rule definition and imposes a constraint. Thus, the rule set can be defined according to the constraint. In addition, rule sets can be defined by using the answers obtained from cyclists through a survey consisting of various questions including their travelling habits. One of the most important advantages of fuzzy logic is that it can easily express linguistic expressions mathematically. Thus, rule sets can be defined with less effort to test the model.

Q.8. "What are the main difficulties that may be encountered in the fuzzy logic model if more factors affecting the route choice are added? How might the addition of more factors affect the setup and accuracy of the fuzzy logic model?"

A fuzzy logic model has been created by considering three different parameters: travel time, traffic safety, and environmental factors. Three different analysis models have been built using binary combinations of these three different parameters, and a fuzzy logic analysis model has been constructed using all of the parameters together. The results obtained by using three different parameters explained the routes followed by the participants of the 'fietstelweek' event more accurately. The increase in the number of parameters taken into account in fuzzy logic models increases the number of rules to be defined. Although increasing the number of rules increases the accuracy of the model, an excessive increase in the number of rules may cause the parameters considered to lose their significance (Trillas & Eciolaza, 2015). In addition, the increasing number of rules requires more time to be spent during the model building and makes the modification of the rule sets more difficult. In order to avoid it, it may be a good way to limit the number of parameters considered and to model some parameters that are considered to be added to the model under other parameters. In addition, the membership degrees of the parameters to be used in the fuzzy logic model developed with ANFIS and the rule relationship can be created automatically by neural network method. ANFIS can be a good alternative to avoid the time spent to manually build the rule systems and constantly modify them in accordance with the outcome.

7. REFERENCES

Akintola, A. A., Aderounmu, G. A., & Owolarafe, O. K. (2002). Problems of engineering education and training in developing countries: Nigeria as a case study. *European Journal of Engineering Education, 27*(4), 393-400. doi:10.1080/03043790210166693

Albertos, P., Sala, A., & Olivares, M. (1998). Fuzzy Logic Controllers. Advantages and Drawbacks. *VII International Congress of Automatic Control. 3*, pp. 833-844. Laxenber, Austria: International Federation o Automatic Control.

Arslan, T., & Khisty, C. (2005). A rational reasoning method from fuzzy perceptions in route choice. *Fuzzy Sets and Systems, 150*(3), 419-435. doi:10.1016/j.fss.2004.03.021

Band, M. (2022). A discrete choice model for route preferences of commuters: comparison between conventional and fast bicycles. *Delft*.

Beheshtitabar, E., Aguilar Ríos, S., König-Hollerwöger, D., Svatý, Z., & Rydergren, C. (2014). ROUTE CHOICE MODELLING FOR BICYCLE TRIPS. *International Journal for Traffic & Transport Engineering, 4*(2), 194-209. doi:10.7708/ijtte.2014.4(2).06

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand.* Boston, MA: MIT Press.

Bicycle Dutch. (2020, August 12). *Bicycle Dutch*. Retrieved from Cycle lanes in the Netherlands: https://bicycledutch.wordpress.com/2020/08/12/cycle-lanes-in-the-netherlands/

Björnerstedt, J., & Verboven, F. (2014). Merger Simulation with Nested Logit Demand. *The Stata Journal: Promoting communications on statistics and Stata, 14*(3), 511-540. doi:10.1177/1536867X1401400304

Bovy, P. H., & Stern, E. (1990). *Route Choice:Wayfinding in Transport Networks.* Dordrecht/The Netherlands: Kluwer Academic Publishers. doi:10.1007/978-94-009-0633-4

Brooks, C. (2008). *RATS Handbook to accompany introductory econometrics for finance.* Cambridge, UK: Cambridge Books.

Bunch, D. S., & Kitamura, R. (1989). Multinomial probit model estimation revisited: testing of new algorithms and evaluation of alternative model specifications for trinomial models of household car ownership. *Institute of Transportation Studies Technical Report, University of California, Davis*.

Bunch, D. S., Louvşere, J. J., & Anderson, D. (1996). A comparison of experimental design strategies for multinomial logit models: The case of generic attributes. *University of California Davis Graduate School of Management Working Paper*, 11-96.

Burrell, J. E. (1968). Multiple route assignment and its application to capacity restraint. *In Proceedings of Fourth International Symposium on the Theory of Traffic Flow.* Karlsruhe/Germany.

Cantillo, V., & Ortúzar, J. D. (2005). A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Transportation Research Part B: Methodological, 39*(7), 641-657. doi:10.1016/j.trb.2004.08.002

Carrasco, J. A., & Ortúzar, J. D. (2002). Review and assessment of the nested logit model. *Transport Reviews, 22*(2), 197-218. doi:10.1080/01441640110091224

Cascetta, E. (2009). Random Utility Theory. In *Transportation Systems Analysis* (pp. 89-167). Boston, MA: Springer.

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. *Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory*, (pp. 697-711). Lyon, France.

Chen, P., & Shen, Q. (2016). Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. *Accident Analysis & Prevention, 86*, 239-246. doi:10.1016/j.aap.2015.11.002

Cheng, S., & Long, J. S. (2007). Testing for IIA in the Multinomial Logit Model. *Sociological Methods & Research, 35*(4), 583-600. doi:10.1177/0049124106292361

Clarry, A., Imani, A. F., & Miller, E. J. (2019). Where we ride faster? Examining cycling speed using smartphone GPS data. *Sustainable Cities and Society, 49*. doi:10.1016/j.scs.2019.101594

Daganzo, C. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forecasting.* New York: Academic Press.

Dhulipala, S., Kedia, A., & Katti, B. (2020). Multi-route choice modelling in a metropolitan context: A comparative analysis using Multinomial Logit and Fuzzy Logic based approaches. *European Transport*. doi:10.48295/ET.2020.79.4

Dial, R. B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research, 5*(2), 83-111. doi:10.1016/0041-1647(71)90012-8

DiGioia, J., Watkins, K., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of Safety Research, 61*, 105-119. doi:10.1016/j.jsr.2017.02.015

Dondi, G., Simone, A., Lantieri, C., & Vignali, V. (2011). Bike Lane Design: the Context Sensitive Approach. *Procedia Engineering, 21*, 897-906. doi:10.1016/j.proeng.2011.11.2092

Dreyfus, G. B., Xu, Y., Shindell, D. T., Zaelke, D., & Ramanathan, V. (2022). Mitigating climate disruption in time: A self-consistent approach for avoiding both near-term and long-term global warming. *Proceedings of the National Academy of Sciences, 119*(22). doi:10.1073/pnas.2123536119

Dubey, S., Mishra, D., Arkatkar, S., Singh, A., & Sarkar, A. (2013). Route choice modelling using fuzzy logic and adaptive neuro-fuzzy. *Modern Traffic and Transportation Engineering Research (MTTER)*, 11-19.

Duckham, M., & Kulik, L. (2003). “Simplest” Paths: Automated Route Selection for Navigation. *International Conference on Spatial Information Theory, 2825*, 169-185. doi:10.1007/978-3-540-39923-0_12

Ehrgott, M., Want, J. Y., Raith, A., & van Houtte, C. (2012). A bi-objective cyclist route choice model. *Transportation Research Part A: Policy and Practice, 46*(4), 652-663. doi:10.1016/j.tra.2011.11.015

Fietstelweek 2016. (n.d.). *Fietstelweek 2016*. Retrieved February 15, 2023, from Results Bicycle Telweek 2016: https://fietstelweek.nl/resultaten-fiets-telweek-bekend/

Fietstelweek. (n.d.). *Fietstelweek*. Retrieved February 15, 2023, from COUNT FOR A BETTER NETHERLANDS CYCLING COUNTRY: https://fietstelweek.nl/

Gärling, T., Kwan, M., & Golledge, R. (1994). Computational-process modelling of household activity scheduling. *Transportation Research Part B: Methodological, 28*(5), 355-364. doi:10.1016/0191-2615(94)90034-5

Gençoğlu, M. T., & Cebeci, M. (1999). Türkiye'de Mühendislik Eğitimi ve Öneriler (Turkish). *Mühendislik-Mimarlık Eğitimi Sempozyumu*, 73-80.

Gommers, M., & Bovy, P. L. (1986). Calibration of the Stochastic Multiple-Route Assignment Model. *Transportation Planning Methods. Proceedings of Seminar M Held at the PTRC Summer Anual Meeting.* Sussex University, England: PTRC Education and Research Services, LTD.

Government of the Netherlands. (n.d.). *Government of the Netherlands/Bicycles*. Retrieved February 12, 2023, from Ways of encouraging bicycle use: https://www.government.nl/topics/bicycles/bicycle-policy-in-the-netherlands

Hawas, Y. (2004). Development and calibration of route choice utility models: neuro-fuzzy approach. *Journal of transportation engineering, 130*(2), 171-182. doi:10.1061/(ASCE)0733-947X(2004)130:2(171)

Henn, V. (2000). Fuzzy route choice model for traffic assignment. *Fuzzy Sets and Systems, 116*(1), 77-101. doi:10.1016/S0165-0114(99)00039-1

Hood, J., Sall, E., & Charlton, B. (2013). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters, 3*(1), 63-75. doi:10.3328/TL.2011.03.01.63-75

Hopkinson, P., & Wardman, M. (1996). Evaluating the demand for new cycle facilities. *Transport Policy, 3*(4), 241-249. doi:10.1016/S0967-070X(96)00020-0

Ivan, J. (2004). New Approach for Including Traffic Volumes in Crash Rate Analysis and Forecasting. *Transportation Research Record: Journal of the Transportation Research Board, 1987*(1), 134-141. doi:0.3141/1897-17

Jones, P. M. (2021). New Approaches to Understanding Travel Behaviour: The Human Activity Approach. In *Behavioural Travel Modelling* (pp. 55-80). Routledge.

Kerr, J., Emond, J. A., Badland, H., Reis, R., Sarmiento, O., Carlson, J., . . . Natarajan, L. (2016). Perceived Neighborhood Environmental Attributes Associated with Walking and Cycling for Transport among Adult Residents of 17 Cities in 12 Countries: The IPEN Study. *Environmental Health Perspectives, 124*(3), 290-298. doi:10.1289/ehp.1409466

Klanjčić, M., Gauvin, L., Tizzoni, M., & Szell, M. (2022). Identifying urban features for vulnerable road user safety in Europe. *EPJ Data Science, 11*(1). doi:10.1140/epjds/s13688-022-00339-5

Koch, T., & Dugundji, E. R. (2021). Taste variation in environmental features of bicycle routes. *IWCTS '21: Proceedings of the 14th ACM SIGSPATIAL International Workshop on Computational Transportation Science*, 1-10. doi:10.1145/3486629.3490697

Koppelman, F. S., & Bhat, C. (2006). A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models. *U.S. Department of Transportation Federal Transit Administration*. Retrieved from https://www.caee.utexas.edu/prof/bhat/COURSES/LM_Draft_060131Final-060630.pdf

Koppelman, F. S., & Wen, C. H. (2000). The paired combinatorial logit model: properties, estimation and application. *Transportation Research Part B: Methodological, 34*(2), 75-89. doi:10.1016/S0191-2615(99)00012-0

Kropko, J. (2007). Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data. *Doctoral dissertation, The University of North Carolina at Chapel Hill.*

Lai, X., & Bierlaire, M. (2015). Specification of the cross-nested logit model with sampling of alternatives for route choice models. *Transportation Research Part B: Methodological, 80*, 220-234. doi:10.1016/j.trb.2015.07.005

Lawrence, B. M., & Oxley, J. A. (2019). You say one route, we observe four: Using naturalistic observation to understand route-choices in cyclists. *Safety Science, 119*, 207-213. doi:10.1016/j.ssci.2019.01.004

Lawson, A. R., Pakrashi, V., Ghosh, B., & Szeto, W. Y. (2013). Perception of safety of cyclists in Dublin City. *Accident Analysis & Prevention, 50*, 499-511. doi:10.1016/j.aap.2012.05.029

Lotan, T., & Koutsopoulos, H. (1993). Route choice in the presence of information using concepts from fuzzy control and approximate reasoning. *Transportation Planning and Technology, 17*(2), 113-126. doi:10.1080/03081069308717505

Lu, W., Scott, D. M., & Dalumpines, R. (2018). Understanding bike share cyclist route choice using GPS data: Comparing dominant routes and shortest paths. *Journal of Transport Geography, 71*, 172-181. doi:10.1016/j.jtrangeo.2018.07.012

Luce, R. (2012). *Individual choice behavior: A theoretical analysis.* Mineola, New York: Dover Publications Inc.

Maddala, G. S. (1986). *Limited-dependent and Qualitative Variables in Econometrics.* Cambridge, UK: Cambridge University Press.

Mamdani, E. H. (1976). Advances in the linguistic synthesis of fuzzy controllers. *International Journal of Man-Machine Studies, 8*(6), 669-678. doi:10.1016/S0020-7373(76)80028-4

Marquart, H., Schlink, U., & Ueberham, M. (2020). The planned and the perceived city: A comparison of cyclists' and decision-makers' views on cycling quality. *Journal of Transport Geography, 82*. doi:10.1016/j.jtrangeo.2019.102602

McFadden, D. (1981). Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272.

Memon, J. A., Demirdöğen, R. E., & Chowdhry, B. S. (2009). Achievements, outcomes and proposal for global accreditation of engineering education in developing countries. *Procedia-Social and Behavioral Sciences, 1*(1), 2557-2561. doi:10.1016/j.sbspro.2009.01.451

Mendel, J. M. (1995). Fuzzy Logic Systems for Engineering: A Tutorial. *Proceedings of the IEEE, 83*(3), 345-377. doi:10.1109/5.364485.

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice, 44*(9), 754-765. doi:10.1016/j.tra.2010.07.008

Moubayed, N., Bernard, M., & Jammal, A. (2009). A survey of engineering education in developing countries: The Lebanese case. *WSEAS Transactions on Advances in Engineering Education, 6*(11), 430-441.

Murat, Y., & Uludag, N. (2008). Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods. *Journal of Scientific & Industrial Research, 67*, 19-27. Retrieved from http://nopr.niscpr.res.in/handle/123456789/731

Oršić, D., Ćosić, M., Šimunović, L., & Sikirić, M. (2022). Geospatial Analysis of Bicycle Infrastructure in the Function of Traffic Safety: The Case of the City of Zagreb. *The Science and Development of Transport-ZIRP 2021*, 15-31. doi:10.1007/978-3-030-97528-9_2

Ortúzar, J., & Willumsen, L. (2011). *Modelling Transport, 4th Edition.* London/UK: John Wiley & Sons, Ltd. doi:10.1002/9781119993308

Osama, A., & Sayed, T. (2017). Evaluating the Impact of Socioeconomics, Land Use, Built Environment, and Road Facility on Cyclist Safety. *Transportation Research Record: Journal of the Transportation Research Board, 2659*(1), 33-42. doi:10.3141/2659-04

Outram, V., & Thompson, E. (1977). Driver route choice. *Proceedings of the PTRC Summer Annual Meeting* (pp. 35-42). Warwick/UK: University of Warwick.

Paleti, R. (2018). Generalized multinomial probit Model: Accommodating constrained random parameters. *Transportation Research Part B: Methodological, 118*, 248-262. doi:10.1016/j.trb.2018.10.019

Prato, C. G., Halldórsdóttir, K., & Nielsen, O. A. (2018). Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space. *International Journal of Sustainable Transportation, 12*(10), 770-781. doi:10.1080/15568318.2018.1437236

Prato, C., & Bekhor, S. (2007). Modeling Route Choice Behavior: How Relevant Is the Composition of Choice Set? *Transportation Research Record: Journal of the Transportation Research Board*, 64-73. doi:10.3141/2003-09

Reynolds, C. C., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental Health, 8*(1), 1-19. doi:10.1186/1476-069X-8-47

Ross, T. J. (2009). *Fuzzy Logic with Engineering Applications.* Singapore: John Wiley & Sons.

Rustum, R., Kurichiyanil, A. M., Forrest, S., Sommariva, C., Adeloye, A. J., Zounemat-Kermani, M., & Scholz, M. (2020). Sustainability ranking of desalination plants using mamdani fuzzy logic inference systems. *Sustainability, 12*(2).

Safiotti, A. (1997). Fuzzy logic in autonomous robotics: behavior coordination. *Proceedings of 6th International Fuzzy Systems Conference* (pp. 573-578). Barcelona, Spain: IEEE. doi:10.1109/FUZZY.1997.616430

Scarf, P., & Grehan, P. (2005). An empirical basis for route choice in cycling. *Journal of Sports Sciences, 23*(9), 919-925. doi:10.1080/02640410400023282

Schepers, J., Kroeze, P., Sweers, W., & Wüst, J. (2011). Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. *Accident Analysis & Prevention, 43*(3), 853-861. doi:10.1016/j.aap.2010.11.005

SciPy. (2023, February 19). *SciPy*. Retrieved March 18, 2023, from Fundamental algorithms for scientific computing in Python: https://scipy.org/

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods.* Englewood Cliffs, New Jersey 07632: Prentice-Hall Inc.

Shi, Q., Abdel-Aty, M., & Lee, J. (2016). A Bayesian ridge regression analysis of congestion's impact on urban expressway safety. *Accident Analysis & Prevention, 88*, 124-137. doi:10.1016/j.aap.2015.12.001

Sobhani, A., Aliabadi, H. A., & Farooq, B. (2019). Metropolis-Hasting based Expanded Path Size Logit model for cyclists' route choice using GPS data. *International Journal of Transportation Science and Technology, 8*(2), 161-175. doi:10.1016/j.ijtst.2018.11.002

Teodorovic, D., & Kikuchi, S. (1990). Transportation route choice model using fuzzy inference technique. *First International Symposium on Uncertainty Modeling and Analysis* (s. 140-145). College Park, MD, USA: IEEE. doi:10.1109/ISUMA.1990.151240

Theocharis, D., Pettit, S., Rodrigues, V. S., & Haider, J. (2018). Arctic shipping: A systematic literature review of comparative studies. *Journal of Transport Geography, 69*, 112-128. doi:10.1016/j.jtrangeo.2018.04.010

Ton, D., Cats, O., Duives, D., & Hoogendoorn, S. (2017). How Do People Cycle in Amsterdam, Netherlands?: Estimating Cyclists' Route Choice Determinants with GPS Data from an Urban Area. *Transportation Research Record: Journal of the Transportation Research Board, 2662*(1), 75-82. doi:10.3141/2662-09

Trillas, E., & Eciolaza, L. (2015). *Fuzzy Logic: An Introductory Course for Engineering Students.* Asturias, Spain: Springer. doi:10.1007/978-3-319-14203-6

Trofimenko, Y., & Shashina, E. (2017). Methodology and Results of Assessing Safety of Bicycle Infrastructure in Russian Cities. *Transportation Research Procedia, 20*, 653-658. doi:10.1016/j.trpro.2017.01.106

Tutorialspoint. (n.d.). *Tutorialspoint - Fuzzy Logic Absolute Beginners*. Retrieved February 17, 2023, from Fuzzy Logic - Set Theory: https://www.tutorialspoint.com/fuzzy_logic/fuzzy_logic_set_theory.htm

Vythoulkas, P., & Koutsopoulos, H. (2003). Modeling discrete choice behavior using concepts from fuzzy set theory, approximate reasoning and neural networks. *Transportation Research Part C: Emerging Technologies, 11*(1), 51-73. doi:10.1016/S0968-090X(02)00021-9

Wall, S., Lee, D., Frangos, S. G., Sethi, M., Heyer, J. H., Ayoung-Chee, P., & Dimaggio, C. J. (2016). The Effect of Sharrows, Painted Bicycle Lanes and Physically Protected Paths on the Severity of Bicycle Injuries Caused by Motor Vehicles. *Safety, 2*(4). doi:10.3390/safety2040026

Wang, C., Quddus, M., & Ison, S. (2013). A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the UK. *Transportmetrica A: Transport Science, 9*(2), 124-148. doi:10.1080/18128602.2010.538871

Wanvik, P. (2009). Effects of road lighting: An analysis based on Dutch accident statistics 1987–2006. *Accident Analysis & Prevention, 41*(1), 123-128. doi:10.1016/j.aap.2008.10.003

Wikimedia Commons. (2009, January 1). *Wikipedia*. Retrieved from Location in Overijssel: https://en.wikipedia.org/wiki/Enschede#/media/File:Map_-_NL_-_Municipality_code_0153_(2009).svg

Wilmsmeier, G., & Martínez-Zarzoso, I. (2009). Determinants of maritime transport costs – a panel data analysis for Latin American trade. *Transportation Planning and Technology, 33*(1), 105-121. doi:10.1080/03081060903429447

Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2010). Built Environment Influences on Healthy Transportation Choices: Bicycling versus Driving. *Journal of Urban Health, 87*, 969-993. doi:10.1007/s11524-010-9509-6

Zadeh, L. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353. doi:10.1016/S0019-9958(65)90241-X

Zhang, Y., & Du, W. (2020). A Simplified C-Logit Stochastic User Equilibrium Model on Bimodal Transportation Network. *Mathematical Problems in Engineering*. doi:10.1155/2020/3702965

8. APPENDICES
Appendix 1

Rule Number	IF	Traffic Safety	AND	Environmental Factors	AND	Travel Time	THEN	Attractiveness (Utility)
Rule 1	IF	Safe	AND	Attractive	AND	Very short	THEN	Very high
Rule 2	IF	Safe	AND	Neutral	AND	Short	THEN	High
Rule 3	IF	Safe	AND	Repellent	AND	Normal	THEN	Medium
Rule 4	IF	Safe	AND	Attractive	AND	Long	THEN	Medium
Rule 5	IF	Safe	AND	Neutral	AND	Very long	THEN	Low
Rule 6	IF	Safe	AND	Repellent	AND	Very short	THEN	High
Rule 7	IF	Safe	AND	Attractive	AND	Short	THEN	Very high
Rule 8	IF	Safe	AND	Neutral	AND	Normal	THEN	High
Rule 9	IF	Safe	AND	Repellent	AND	Long	THEN	Low
Rule 10	IF	Safe	AND	Attractive	AND	Very long	THEN	Low
Rule 11	IF	Safe	AND	Neutral	AND	Very short	THEN	Very high
Rule 12	IF	Safe	AND	Repellent	AND	Short	THEN	High
Rule 13	IF	Safe	AND	Attractive	AND	Normal	THEN	High
Rule 14	IF	Safe	AND	Neutral	AND	Long	THEN	Medium
Rule 15	IF	Safe	AND	Repellent	AND	Very long	THEN	Low
Rule 16	IF	Normal	AND	Attractive	AND	Very short	THEN	Very high
Rule 17	IF	Normal	AND	Neutral	AND	Short	THEN	High
Rule 18	IF	Normal	AND	Repellent	AND	Normal	THEN	Medium
Rule 19	IF	Normal	AND	Attractive	AND	Long	THEN	Medium
Rule 20	IF	Normal	AND	Neutral	AND	Very long	THEN	Low
Rule 21	IF	Normal	AND	Repellent	AND	Very short	THEN	High
Rule 22	IF	Normal	AND	Attractive	AND	Short	THEN	High
Rule 23	IF	Normal	AND	Neutral	AND	Normal	THEN	Medium
Rule 24	IF	Normal	AND	Repellent	AND	Long	THEN	Low
Rule 25	IF	Normal	AND	Attractive	AND	Very long	THEN	Low
Rule 26	IF	Normal	AND	Neutral	AND	Very short	THEN	High
Rule 27	IF	Normal	AND	Repellent	AND	Short	THEN	Medium
Rule 28	IF	Normal	AND	Attractive	AND	Normal	THEN	High
Rule 29	IF	Normal	AND	Neutral	AND	Long	THEN	Medium
Rule 30	IF	Normal	AND	Repellent	AND	Very long	THEN	Very low
Rule 31	IF	Unsafe	AND	Attractive	AND	Very short	THEN	High
Rule 32	IF	Unsafe	AND	Neutral	AND	Short	THEN	Medium
Rule 33	IF	Unsafe	AND	Repellent	AND	Normal	THEN	Low
Rule 34	IF	Unsafe	AND	Attractive	AND	Long	THEN	Medium
Rule 35	IF	Unsafe	AND	Neutral	AND	Very long	THEN	Very low
Rule 36	IF	Unsafe	AND	Repellent	AND	Very short	THEN	Medium
Rule 37	IF	Unsafe	AND	Attractive	AND	Short	THEN	High
Rule 38	IF	Unsafe	AND	Neutral	AND	Normal	THEN	Medium
Rule 39	IF	Unsafe	AND	Repellent	AND	Long	THEN	Very low
Rule 40	IF	Unsafe	AND	Attractive	AND	Very long	THEN	Low
Rule 41	IF	Unsafe	AND	Neutral	AND	Very short	THEN	Medium
Rule 42	IF	Unsafe	AND	Repellent	AND	Short	THEN	Medium
Rule 43	IF	Unsafe	AND	Attractive	AND	Normal	THEN	Low
Rule 44	IF	Unsafe	AND	Neutral	AND	Long	THEN	Low
Rule 45	IF	Unsafe	AND	Repellent	AND	Very long	THEN	Very low

Appendix 2

```
import numpy as np
import skfuzzy as fuzz
from skfuzzy import control as ctrl
import pandas as pd
import os

#Traffic Safety

traffic_safety = ctrl.Antecedent(np.arange(0, 61, 1), 'Traffic Safety')

traffic_safety['Safe'] = fuzz.trimf(traffic_safety.universe, [30, 60, 60])
traffic_safety['Normal'] = fuzz.trimf(traffic_safety.universe, [15, 30, 45])
traffic_safety['Unsafe'] = fuzz.trimf(traffic_safety.universe, [0, 0, 30])

traffic_safety.view()

#Environmental factors

environmental = ctrl.Antecedent(np.arange(0, 51, 1), 'Environmental Factors')

environmental['Attractive'] = fuzz.trimf(environmental.universe, [25, 50, 50])
environmental['Neutral'] = fuzz.trimf(environmental.universe, [15, 25, 35])
environmental['Repellent'] = fuzz.trimf(environmental.universe, [0, 0, 25])

environmental.view()

#Travel time

travel_time = ctrl.Antecedent(np.arange(0, 101, 1), 'Travel Time')

travel_time['Very short'] = fuzz.trimf(travel_time.universe, [75, 100, 100])
travel_time['Short'] = fuzz.trimf(travel_time.universe, [60, 75, 90])
travel_time['Normal'] = fuzz.trimf(travel_time.universe, [40, 55, 70])
travel_time['Long'] = fuzz.trimf(travel_time.universe, [20, 40, 60])
travel_time['Very long'] = fuzz.trimf(travel_time.universe, [0, 0, 40])

travel_time.view()

#Attractiveness

attractive = ctrl.Consequent(np.arange(0, 101, 1), 'Attractiveness')

attractive['Very high'] = fuzz.trimf(attractive.universe, [80, 100, 100])
attractive['High'] = fuzz.trimf(attractive.universe, [60, 75, 90])
attractive['Medium'] = fuzz.trimf(attractive.universe, [45, 60, 75])
attractive['Low'] = fuzz.trimf(attractive.universe, [25, 45, 65])
attractive['Very low'] = fuzz.trimf(attractive.universe, [0, 0, 45])

attractive.view()

# Rule set when traffic safety is safe

rule1 = ctrl.Rule(traffic_safety['Safe'] & environmental['Attractive'] &
travel_time['Very short'], attractive['Very high'])
```

```
rule2 = ctrl.Rule(traffic_safety['Safe'] & environmental['Neutral'] &
travel_time['Short'], attractive['High'])
rule3 = ctrl.Rule(traffic_safety['Safe'] & environmental['Repellent'] &
travel_time['Normal'], attractive['Medium'])
rule4 = ctrl.Rule(traffic_safety['Safe'] & environmental['Attractive'] &
travel_time['Long'], attractive['Medium'])
rule5 = ctrl.Rule(traffic_safety['Safe'] & environmental['Neutral'] &
travel_time['Very long'], attractive['Low'])
rule6 = ctrl.Rule(traffic_safety['Safe'] & environmental['Repellent'] &
travel_time['Very short'], attractive['High'])
rule7 = ctrl.Rule(traffic_safety['Safe'] & environmental['Attractive'] &
travel_time['Short'], attractive['Very high'])
rule8 = ctrl.Rule(traffic_safety['Safe'] & environmental['Neutral'] &
travel_time['Normal'], attractive['High'])
rule9 = ctrl.Rule(traffic_safety['Safe'] & environmental['Repellent'] &
travel_time['Long'], attractive['Low'])
rule10 = ctrl.Rule(traffic_safety['Safe'] & environmental['Attractive'] &
travel_time['Very long'], attractive['Low'])
rule11 = ctrl.Rule(traffic_safety['Safe'] & environmental['Neutral'] &
travel_time['Very short'], attractive['Very high'])
rule12 = ctrl.Rule(traffic_safety['Safe'] & environmental['Repellent'] &
travel_time['Short'], attractive['High'])
rule13 = ctrl.Rule(traffic_safety['Safe'] & environmental['Attractive'] &
travel_time['Normal'], attractive['High'])
rule14 = ctrl.Rule(traffic_safety['Safe'] & environmental['Neutral'] &
travel_time['Long'], attractive['Medium'])
rule15 = ctrl.Rule(traffic_safety['Safe'] & environmental['Repellent'] &
travel_time['Very long'], attractive['Low'])

# Rule set when traffic safety is normal

rule16 = ctrl.Rule(traffic_safety['Normal'] & environmental['Attractive'] &
travel_time['Very short'], attractive['Very high'])
rule17 = ctrl.Rule(traffic_safety['Normal'] & environmental['Neutral'] &
travel_time['Short'], attractive['High'])
rule18 = ctrl.Rule(traffic_safety['Normal'] & environmental['Repellent'] &
travel_time['Normal'], attractive['Medium'])
rule19 = ctrl.Rule(traffic_safety['Normal'] & environmental['Attractive'] &
travel_time['Long'], attractive['Medium'])
rule20 = ctrl.Rule(traffic_safety['Normal'] & environmental['Neutral'] &
travel_time['Very long'], attractive['Low'])
rule21 = ctrl.Rule(traffic_safety['Normal'] & environmental['Repellent'] &
travel_time['Very short'], attractive['High'])
rule22 = ctrl.Rule(traffic_safety['Normal'] & environmental['Attractive'] &
travel_time['Short'], attractive['High'])
rule23 = ctrl.Rule(traffic_safety['Normal'] & environmental['Neutral'] &
travel_time['Normal'], attractive['Medium'])
rule24 = ctrl.Rule(traffic_safety['Normal'] & environmental['Repellent'] &
travel_time['Long'], attractive['Low'])
rule25 = ctrl.Rule(traffic_safety['Normal'] & environmental['Attractive'] &
travel_time['Very long'], attractive['Low'])
rule26 = ctrl.Rule(traffic_safety['Normal'] & environmental['Neutral'] &
travel_time['Very short'], attractive['High'])
rule27 = ctrl.Rule(traffic_safety['Normal'] & environmental['Repellent'] &
travel_time['Short'], attractive['Medium'])
rule28 = ctrl.Rule(traffic_safety['Normal'] & environmental['Attractive'] &
travel_time['Normal'], attractive['High'])
```

```
rule29 = ctrl.Rule(traffic_safety['Normal'] & environmental['Neutral'] &
travel_time['Long'], attractive['Medium'])
rule30 = ctrl.Rule(traffic_safety['Normal'] & environmental['Repellent'] &
travel_time['Very long'], attractive['Very low'])

# Rule set when traffic safety is unsafe

rule31 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Attractive'] &
travel_time['Very short'], attractive['High'])
rule32 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Neutral'] &
travel_time['Short'], attractive['Medium'])
rule33 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Repellent'] &
travel_time['Normal'], attractive['Low'])
rule34 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Attractive'] &
travel_time['Long'], attractive['Medium'])
rule35 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Neutral'] &
travel_time['Very long'], attractive['Very low'])
rule36 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Repellent'] &
travel_time['Very short'], attractive['Medium'])
rule37 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Attractive'] &
travel_time['Short'], attractive['High'])
rule38 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Neutral'] &
travel_time['Normal'], attractive['Medium'])
rule39 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Repellent'] &
travel_time['Long'], attractive['Very low'])
rule40 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Attractive'] &
travel_time['Very long'], attractive['Low'])
rule41 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Neutral'] &
travel_time['Very short'], attractive['Medium'])
rule42 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Repellent'] &
travel_time['Short'], attractive['Medium'])
rule43 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Attractive'] &
travel_time['Normal'], attractive['Low'])
rule44 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Neutral'] &
travel_time['Long'], attractive['Low'])
rule45 = ctrl.Rule(traffic_safety['Unsafe'] & environmental['Repellent'] &
travel_time['Very long'], attractive['Very low'])


attractiveness_ctrl = ctrl.ControlSystem([rule1, rule2, rule3, rule4, rule5, rule6,
rule7, rule8, rule9, rule10, rule11, rule12, rule13, rule14, rule15, rule16, rule16,
rule17, rule18, rule19, rule20, rule21, rule22, rule23, rule24, rule25, rule26,
rule27, rule28, rule29, rule30, rule31, rule32, rule33, rule34, rule35, rule36,
rule37, rule38, rule39, rule40, rule41, rule42, rule43, rule44, rule45])
attractiveness_result = ctrl.ControlSystemSimulation(attractiveness_ctrl)




# Read the Excel file
df = pd.read_excel('Excel file name.xlsx', sheet_name='Sheet Name')

# Get the values from the columns
traffic_safety = df['Traffic Safety'].tolist()
environmental_factors = df['Environmental Factors'].tolist()
travel_time = df['Travel Time'].tolist()
```

```python
# Combine the values into a list of tuples
values = list(zip(traffic_safety, environmental_factors, travel_time))

attractiveness_values = []
for i, (traffic_safety, environmental_factors, travel_time) in enumerate(values):
    attractiveness_result.input['Traffic Safety'] = traffic_safety
    attractiveness_result.input['Environmental Factors'] = environmental_factors
    attractiveness_result.input['Travel Time'] = travel_time
    attractiveness_result.compute()
    attractiveness_values.append(attractiveness_result.output['Attractiveness'])

df['Attractiveness of the Route'] = attractiveness_values

df.to_excel(r"Folder location", index=False)
```