Doctoral Dissertation (Shinshu University)

Evolutionary design of sustainable mobility and transport system

March 2018

Armas Andrade Tito Rolando

SHINSHU UNIVERSITY

Abstract

Department of Mathematics and System Development
Interdisciplinary Graduate School of Science and Technology

Doctor of Philosophy

Evolutionary Design of Sustainable Mobility and Transport System

by Rolando ARMAS

Cities are highly relevant since the last decades. They are common areas where people, meet, live and perform several activities. Mobility is the way to move or be moved freely and easily and includes infrastructure and services demand. For the managers and decision makers of the city, to guarantee an acceptable mobility level of service is a continuous challenge because the mobility and transportation system involve huge geographical areas, many variables, components and interactions among them and therefore shifts in a complex problem. Besides, developing a sustainable system where social, economic and environmental aspects are taken into account, imposes additional difficulties and constraints. For a decision maker, dealing with all of those aspects variables and relationships is troublesome because the variable space turns vast and some objectives can be in conflict with each other.

This work focuses on a framework implementation that joins evolutionary computing, mobility and transportation simulators and data mining techniques. The aim is to provide a method and tools that allow the exploration of the potential solutions and also the trajectory of the evolution. The framework starts with mobility problem modelling according to a specific case of study. After that, evolutionary computing with single and multi-many objective optimization evolutionary algorithms is used to explore the tentative solutions according to the model. The outcome of optimization is analyzed in the variable and objective space to identify patterns, understand better the model, and extract some knowledge for the decision maker. The framework is iterative, where the knowledge extracted for the decision maker can be used as feedback to fine tune the model or to study more complex formulations of the problem.

In this work, Quito city's business centre is used as real world case study which covers approximately 5x8 km^2. This work implements a mobility model for Quito, including mobility plans and the transportation network. Also, it studies traffic signal optimization formulating the problem from a single and multi-objective optimization perspective. Additionally, it studies level of service in public and private transportation using three-objective and bi-objective formulations of the problem.

As a first scenario, a 70 traffic signals optimization scenario is executed modelling the mobility of 20.000 agents that use private transportation. It is a large scenario not only due to the geographical area but also the variable space size which implies an expensive simulation computational time to evaluate each solution. A set of genetic operators is proposed to accelerate the convergence of the algorithm and promote a better solution's configuration for inducing continuous traffic flow among intersection neighbours. Several experiments perform single objective approach to minimize travel time.

A complete analysis is given in three type of spaces: the variable, the objective and the geographical. In the variable space, analysis verifies that the neighbourhood operators induce coordination between signals and find some patterns which are identified using clustering methods. The identified groups are geolocated, verifying that they are positioned in neighbourhoods as expected. Such clusters of coordinated signals favour the continuous traffic flows. In the objective space, analysis verifies that the groups of coordinated signals no only benefit the reduction of travel time, but also decline emissions. Analyzing the difference in CO2 emissions between two solutions, one in the initial population and the other an optimized one, emissions reductions across the area of study when signals are optimized is found.

In another study, a bi-objective optimization approach to minimize travel time and fuel consumption simultaneously is conducted. A set of experiments using the same mobility scenario but changing the emission model is performed. Mainly, the study examines the conflict between objectives, if any, when they are optimized simultaneously and how the settings of the signals relate to the trade-offs between them. One of the challenges in this study is the computational time to evaluate the emissions and fuel consumption. That imposes an investigation to figure out a method to accelerate the evolution using small populations with relative small evaluations. In part, it is achieved by the implementation of deterministic varying mutation operator used in single-objective. Results show that the optimization of signals enabling different cycle times and coordinating them by adequately setting their offsets, lead to significant reductions in both fuel consumption and travel time. The small number of non-dominated solutions in the last generation show that both objective functions are correlated.

In this research, a different problem examines traffic density levels in urban transportation. The study contemplates the urban transportation system under various proportions of private and public transportation users. The purpose is to explain the conditions to achieve different levels of service and their relationships with the optimal configuration of the public transportation, traffic density, travel time, fuel consumption and emissions. The level of service (LoS) refers to quality measures in a traffic stream. In this work, LoS is defined concerning to traffic density. The scenario is based in Quito's business center, modelling the mobility of 27.000 agents that use private and public transportation. In Quito, the massive passenger transportation is operated by private and public companies. Five main Bus Rapid Transit (BRT) corridors which are the most demanded and congested routes are considered for this mobility model. The levels of service are influenced by the proportions of the population that uses public and private transportation and translates to a bi-level optimization problem. For a given proportion, the model determines a proper configuration of public transportation system in terms of capacities of the buses and departure times between buses. A first scenario concerns with three-objective optimization, focus on minimizing travel time, fuel consumption and traffic density. The Adaptive ϵ-Sampling and ϵ-Hood (AϵSϵH) algorithm is adopted to search the optimal solutions. A set of experiments and a trade-off analysis between objectives is conducted. The negative correlation between objectives verified a trade-off between travel time versus fuel consumption and travel time versus density. The results show that in general solutions with the best LoS, have low values of fuel consumption, but high travel times in opposite with solutions with worst levels of service that show best travel time but the worst fuel consumption. Sustainable public transportation implies a low environmental impact. In general, the results prove that best levels of service and low levels of fuel consumption can be achieved simultaneously. An analysis in Particulate Matter (PM) emission and its geo-location is conducted between two optimal solutions from the Pareto optimal set. The analysis pointed that a reduction in PM depends on BRT headways and technology (fuel type).

Another study concern to analyze the same scenario as a bi-objective optimization problem, focusing only on travel time and traffic density as optimization objectives. Comparable results show a confirmation that better LoS (low density) corresponds to a high proportion of public transportation users. Also, a new set of solutions with the same LoS is chosen to make a complete analysis of the variable, objective and geographical space. This complementary analysis revealed that even if the solutions share the same LoS with different configurations in capacities and headways, effects in pollution must be taken into consideration by the decision makers when they decide to improve the transportation system.

Finally, this work is summarized presenting the conclusions and future work.

Acknowledgements

I want to express my special gratitude to my main advisor Prof. Hernan Aguirre and my supervisor Prof. Kiyoshi Tanaka. To the thesis committee members: Prof. Toshiro Sato, Prof. Hiroyuki Sato, and Prof. Masayuki Nakamura. Thanks to the Faculty members at Shinshu University for their support and advice. I gratefully acknowledge the financial support provided by Ecuador Government through the support of National Secretariat of Higher Education, Science, Technology and Innovation. Special thanks to all my friends. Finally, thanks to my family for their support and help.

Contents

Abstract iii

Acknowledgements vii

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Motivation & Goal . . . 3
- 1.3 Contribution . . . 3
- 1.4 Outline . . . 4

2 Design Optimization Framework 7
- 2.1 Introduction . . . 7
- 2.2 Components . . . 7
- 2.3 Problem Formulation . . . 8
- 2.4 Simulation . . . 9
  - 2.4.1 Multi-Agent Transport Simulation MATSim . . . 9
  - 2.4.2 Modules . . . 11
    - Emissions . . . 11
    - Traffic Signals . . . 11
  - 2.4.3 Comprehensive Modal Emission Model (CMEM) . . . 12
- 2.5 Evolutionary Algorithms . . . 12
  - 2.5.1 Main Concepts and Definitions . . . 12
  - 2.5.2 Single Objective Optimization . . . 13
  - 2.5.3 Multi-Many Objective Optimization . . . 14
  - 2.5.4 EA used in this work . . . 15
    - Elitism Single Objective EA . . . 15
    - MOEA:Adaptive ε-Sampling and ε-Hood AεSεH Algorithm . . . 17
    - MOEA: Adaptive $\boldsymbol{\varepsilon}$-Sampling and $\boldsymbol{v}$-Hood (A$\boldsymbol{\varepsilon S v H}$) Algorithm . . . 18
- 2.6 Analysis of Solutions . . . 21
- 2.7 Design Knowledge . . . 21

3 Quito Model Mobility Scenario 23
- 3.1 Quito Metropolitan District-DMQ . . . 23
- 3.2 Area of Study . . . 25
- 3.3 How to build DMQScenario . . . 26
  - 3.3.1 Transport Network Infrastructure . . . 26
  - 3.3.2 Initial Demand - Mobility Model . . . 28
- 3.4 Evolutionary Algorithm and MATSim Integration . . . 31

3.4.1 AnyOEA Framework . . . 31
3.4.2 MATSim Process . . . 34
3.4.3 AnyOEA-MATSim Integration . . . 34

4 Traffic Signals Optimization - Single Objective Formulation 37
4.1 Introduction . . . 37
4.2 Related Work . . . 38
4.3 Problem Formulation . . . 40
4.4 Simulation and EA Integration . . . 42
4.5 Evolutionary Algorithm . . . 42
4.5.1 Traffic Signals Representation . . . 42
4.5.2 Operators of Variation . . . 43
4.5.3 Varying mutation schedule . . . 44
4.5.4 Fitness Function . . . 44
4.6 Experimental Setup . . . 45
4.6.1 Evolutionary Algorithm . . . 45
4.6.2 Initial Mobility Plans . . . 45
4.7 Simulation Results and Discussion . . . 46
4.7.1 Effects of Operators . . . 46
4.7.2 Algorithm's Parameters Analysis . . . 51
4.7.3 Vehicle Movement Distribution . . . 52
4.7.4 Analysis of Solutions:Decision Space and Cluster Analysis . . . 53
4.7.5 Emissions and Fuel Consumption . . . 57
4.7.6 Spatial Analysis for CO_2 Emissions . . . 59
4.7.7 Coordination Analysis . . . 60
4.8 Conclusions and Future Work . . . 62

5 Traffic Signals Optimization - Bi Objective Formulation 65
5.1 Introduction . . . 65
5.2 Multi-objectivization . . . 66
5.3 Evolutionary Algorithm . . . 67
5.3.1 Representation & Operators . . . 67
5.3.2 Fitness Functions . . . 67
5.4 MATSim and CMEM Preliminaries . . . 68
5.5 Evolutionary Algorithm Experimental Setup . . . 68
5.6 Simulation Results and Discussion . . . 69
5.7 Conclusion and Future Work . . . 73

6 Level of Service Optimization: Bi-Objective Formulation 75
6.1 Introduction . . . 75
6.2 Method . . . 77
6.2.1 Components & Overview . . . 77
6.2.2 Problem Definition . . . 78
6.2.3 Area of Study . . . 78
6.2.4 Simulation and MOEA Integration . . . 79
6.3 Evolutionary Algorithm . . . 80
6.3.1 Solution Representation . . . 80

6.3.2 Operators . . . 81
6.4 Evaluation Functions . . . 82
6.4.1 Travel Time f_1 . . . 82
6.4.2 Density f_2 . . . 82
6.5 Simulation Results and Discussion . . . 83
6.5.1 Experimental Setup . . . 83
6.5.2 Comparison between AεSεH and AεSvH . . . 84
6.5.3 Trade-offs . . . 86
6.5.4 Emissions . . . 89
6.6 Conclusions and Future Work . . . 90

7 Level of Service: Three-Objective Formulation **95**
7.1 Introduction . . . 95
7.2 Evaluation Functions . . . 95
7.2.1 Travel Time . . . 95
7.2.2 Density . . . 96
7.2.3 Fuel Consumption & Emissions . . . 96
7.3 Simulation Results and Discussion . . . 96
7.3.1 Hypervolume . . . 96
7.3.2 Trade-offs . . . 97
7.3.3 Emissions . . . 98
7.4 Conclusions and Future Work . . . 100

8 Conclusions **105**

Bibliography **109**

List of Figures

2.1 Method . . . 8
2.2 Matsim process . . . 10
2.3 Traffic Light MATSim Representation . . . 12
2.4 Evolutionary Algorithm Flow Chart . . . 16

3.1 DMQ Urban Area (captured by google earth) . . . 23
3.2 Ecuador (1968-2008) Inhabitants vs. Vehicles . . . 24
3.3 Area of Study . . . 25
3.4 Base DMQ Scenario Implementation . . . 26
3.5 DMQ from OpenStreetMap . . . 27
3.6 DMQ Scenario Network . . . 27
3.7 Saturation Flow and Network Storage . . . 28
3.8 DMQ Scenario Demand Modeling . . . 28
3.9 Origin and Destination Locations . . . 30
3.10 Evolutionary Algorithm Framework . . . 33
3.11 AnyOEA-MATSim Integration . . . 35

4.1 Traffic light components . . . 41
4.2 Simulation and EA integration . . . 43
4.3 Solution representation . . . 43
4.4 Cycle and offset propagation . . . 44
4.5 Mobility Scenarios S124h, S2M and S2A . . . 46
4.6 Mean travel time over generations in experiments E3,E4 and E4DVM . . . 48
4.7 Mean travel time of best solutions in experiments E3,E4 and E4DVM . . . 48
4.8 Expected number of mutated signals with constant and varying mutation schedules (E4,E4DVM) . . . 48
4.9 Mean travel time over generations, scenarios S2M and S2A (E4, E4DVM) . . . 49
4.10 Effect of crossover operator on mean travel time . . . 50
4.11 DMQ scenario flow from 07:30AM to 08:30AM . . . 53
4.12 Signals Corridor S1-S15 . . . 53
4.13 Signals Corridor S3-S5 flow from 07:30AM to 08:30AM . . . 54
4.14 Cycle length of the best solutions for experiment E4DVM scenario S2M. Annotated heatmap: solutions by row (associated travel time in the row label), variables by column (one per signal), values by color (see color legend). . . . 54
4.15 Offset times of the best solutions for experiment E4DVM scenario S2M . . . 55

4.16 Hierarchical clustering of cycle length of the best solutions for experiment E4DVM scenario S2M 55
4.17 Geolocation of signal clusters, experiment E4DVM scenario S2M 56
4.18 Geolocation of signal clusters, experiment E4DVM scenario S124h 57
4.19 Change in CO_2 emissions 60
4.20 Variables to detect green phase crossing 61
4.21 Best individual (lowest travel time) cluster No.6 (a) and cluster No. 9 (b) for South-North direction, experiment E4DVM, scenario S2M 62

5.1 Optimization System 66
5.2 Objective values of solutions by bi- and single-objective optimization 69
5.3 Non-dominated solutions in initial population labeled with cycle length 70
5.4 Cycle length, best travel time solutions, single- and bi-objective optimization 70
5.5 Best Solution Cycle 72
5.6 Best Solution Offset 72
5.7 Traffic Volume 74

6.1 DMQ BRT Scenario 79
6.2 Simulation and EA integration 80
6.3 Solution Representation 81
6.4 Hypervolume Over Generations 85
6.5 Hypervolume Over Generations of One Run. 85
6.6 Pareto Optimal Set of one run. 86
6.7 Travel time and density, colored by fraction of public transport users n_{Pt}. Black dots show POS founded by the original algorithm. 88
6.8 Fuel consumption and travel time, colored by level of service LoS. 88
6.9 Fuel consumption and density, colored by fraction of public transport users n_{Pt}. 89
6.10 Trade-off TT and PM by n_{Pt} 91
6.11 Number of Public Transportation Trips by n_{Pt} 91
6.12 Public Transportation PM Emission by n_{Pt} 91
6.13 Configuration of bus Capacities (top) and Headways (bottom) of solutions A8b and A8a. Schedules: Mo, Mi, A, N (morning, midday, afternoon, night). Capacities: 0, 1, 2 (small, medium, large). Headways: 0, $\cdots$, 11 (5 min,$\cdots$, 60 min). 92
6.14 PM Emission Difference Between A8b and A8a 92

7.1 Hyper Volume Over Generations 97
7.2 Fuel consumption and travel time, colored by level of service LoS 99

7.3 Fuel consumption and travel time, colored by fraction of public transport users N_{Pt} 99
7.4 Travel time and density, colored by fraction of public transport users N_{Pt} 99
7.5 Trade-off TT and PM (By LoS) 101
7.6 Number of Public Transportation Trips by LoS 101
7.7 Public Transportation PM Emission by LoS 101
7.8 POS Levels A and B (by N_{Pt}) 102
7.9 Travel Time vs PMPT Trade-Off 102
7.10 PM Emission Difference Between A7 and B6 103

List of Tables

3.1 Reasons for trips in DMQ 24
3.2 Scenario Inhabitants per District (2010) 29
3.3 Notation in mobility scenario 31
3.4 DMQ Mobility reasons 31

4.1 Probability of mutation operators P_{CtO} and P_{Gt} and mutation probability P_m and $P_m^{(t)}$ per signal. 45
4.2 Mean travel time, standar deviation and interquartile range of best solutions in experiment E3,E4 and E4DVM. 47
4.3 SMAC solutions: EA parameters 51
4.4 Expected Number of Mutated Signals by SMAC configuration and E4DVM 51
4.5 Car distribution (fuel=gasoline and weight $\leqq$ 2Tons) 58
4.6 Emissions of best solutions on scenario S124h 59
4.7 Emissions of best solutions on scenario S2M. 59
4.8 Green waves cluster 6 total agents=253 62
4.9 Green waves cluster 9 total agents=1400 62

5.1 CMEM Vehicle Categorization 68
5.2 Scenario Emissions 73

6.1 **Automovile distribution (fuel=gasoline and weight $\leq$ 2Tons)** 84
6.2 Two Set Coverage Index (C) 86
6.3 Objective Correlation Matrix 86
6.4 Levels of service for basic freeways segments 87

7.1 Objective Correlation Matrix 97
7.2 Levels of service for basic freeways segments 97

Dedicated to my Mom

Chapter 1

Introduction

This chapter presents an overview of the research. The background, motivation and the objective of the study are introduced together with a list and description of the contributions. At the end of the chapter, the outline of this work is presented.

1.1 Background

Transportation exists at the foundation of any functioning city, providing mobility and access to jobs, amenities and resources. As a result, transportation systems are a vital component of building sustainable cities and have a profound impact on both the global and local environment.
Transportation and mobility systems have become large-scale and complex in many cities, with broad social impacts and strong implications for the economy and the environment. Urban population is overgrowing around the globe (United Nations and Social Affairs, 2014) bringing an increase in transportation and mobility demand that improperly satisfied often causes congestion of the system.

Traffic congestion has become a problem in different cities around the world. Also, road transportation is one of the main human-made greenhouse gas sources. Mobility related pollution shares around 11% of global greenhouse gas emissions. Also, congestion adds substantial costs due to delays, the risk of accidents and health concern. Thus, efforts to continuously redesign the transportation system and make it sustainable to guarantee the mobility of larger populations and the accessibility of cities are required. These efforts include the increase of infrastructure capacity, land-use planning, improving public transportation, and the incorporation of soft-computing methods to implement Intelligent Transportation Systems (ITS).

In big cities located around the world, the effects of traffic congestion due to increment in citizen trips to accomplish diary tasks such as work, study, shopping or leisure, is a testimony that traffic jams (Sugiyama et al., 2008) become a serious issue for residents living in that cities. Also in urban areas located in developing countries, the travel times, usually are high because the number of vehicles is continuously increasing (Gakenheimer, 1999) while the infrastructure is not advancing at the same pace to support the demand,

producing congestion and affecting all aspects of live.

To gain a better understanding of complex sustainable systems and propose solutions to their problems is fundamental first to model the current systems. Currently, there are various simulators available that can be used by transport engineers to define scenarios, analyse, and understand transportation and mobility problems that arise in cities. However, the complexity of a model that captures a whole city could be overwhelming even for an expert. So, the expert usually limits the scope of its exploration to small local areas and intersections of the city and can perform simulations only on few alternative scenarios. Thus, the space of alternatives an expert observe is quite limited and incomplete, even by using sophisticated transport and mobility simulators.

Complex problems usually are characterized by properties such as multi-many objectives formulations, multi-modality, nonlinearity, discreteness, hard and soft constraints and non-separability(Reed et al., 2013). Thus, meta-heuristics like multi-objective evolutionary algorithms (MOEA) play a major role in dealing with that kind of problems. MOEAs provide the means to search and explore several alternatives, allowing the human expert to learn about the solution space and focus on the most suitable solutions. MOEAs are population-based iterative processes that need several iterations to achieve good results. However, evaluating mobility scenario is very costly, since it requires the generation of a mobility plan for a synthetic population of agents, and the simulation of such plan in an agent-based model. This seriously limits the population size and the number of generations that the evolutionary algorithm can use. In addition, to get a better understanding of the problems of the city and provide global solutions to them, we should simulate a significant part of the traffic in the city, which implies a large search space. Thus, the problem is at the same time computationally expensive and large-scale.

This work focuses on Quito city (Ecuador) as an example of a real world case study. Quito is a large city where almost two millions of persons commute every day for several reasons as work, study, leisure, shopping etc., using private and public transportation. The area of study includes the business district, eight major universities, several hospitals, large malls, two large parks, and one major soccer stadium, covering approximately 5×8 Km^2. The varied and large number of services and the abundant population make the existing traffic infrastructure clearly not well engineered for the daily demand. Quito's urban population growth has put enormous strains on its transportation system. There is the need to model the system, better understand the problems associated with transportation and mobility, find sustainable solutions to solve them and improve life in the city.

1.2 Motivation & Goal

The research aims to implement a framework that integrates evolutionary computation, traffic simulation, emission models, and data mining tools. That framework will facilitate the transition from a current transportation and mobility system to a desired optimal one, bridging the gap between policy making and the required design and implementation of the new system that secures the sustainability of the city.
The proposed framework will be of great value to transport and mobility engineers as well as city planners and decision makers. It will allow to the expert leave the exploration of a vast number of possible designs and scenarios to intelligent optimisation tools and concentrate on directing evolution and analysing promising designs and scenarios found in artificial evolution. The objective of this research is to develop a framework that integrates transport and mobility simulators with emission models, evolutionary optimizers, and data mining tools. Implement the methods to improve the performance of EA for single and multi-objective optimization for design innovation of large-scale transport and mobility systems. As a proof of concept, we choose Quito Metropolitan District (DMQ), Capital of Ecuador, as a test case. Supported by data mining tools, we analyse Pareto optimal designs to extract trade-off knowledge.

1.3 Contribution

Research contributions of this work are as follows:

A framework for testing large-scale evolutionary design problem's hypothesis related to mobility and transportation: The method implemented provides a guide, methods, tools to formulate and test hypothesis related to mobility and transportation problems. Based on evolutionary design optimization and coupling single and multi-many objective evolutionary algorithms with traffic and emission simulators; the framework implemented is a virtual laboratory for run experiments with complex models.

The framework developed is a powerful tool to explore and analyse optimized solutions: A complete method which explores three different spaces: variable, objective and geographical have been implemented. Those three spaces are interrelated; the framework implemented concede a complete and comprehensive analysis from that three different perspectives.

Operators that run on two-dimensional neighbourhoods of traffic signals aiming to quickly find clusters of coordinated signals by propagating cycle times and setting offsets based on the distance between signals. Intuitively, we expect groups of correlated signals for which similar settings could be appropriate for given scenarios. The algorithm searches for optimal combinations of correlated signals. Due to the topography of the city and its mobility patterns, it is important to define neighbourhoods as two-dimensional and optimize in both directions of traffic flow (north-south-north

and east-west-east). These two-dimensional correlations increase the complexity of the problem substantially concerning to optimization.

Introduces varying mutation rates with high selection pressure to accelerate the algorithm convergence: The appropriate combination of operators, selection, and varying mutation rates allows searching effectively even with small populations. The problem is computationally expensive, for that reason, a rapid convergence is highly relevant. Moreover, a quick convergence and fewer evaluations (simulations) benefit to the EA to be scalable.

In addition, the work presents a detailed **analysis of the algorithm design and its parameters** showing alternative ways the algorithm can be used to find solutions to the given problem, which is valuable for the non-expert in evolutionary algorithms.

Improvement in the analysis and understanding of the level of service in urban transportation: The model and the framework provide a set of outcomes to study and learn the effects of changing the proportion of cars vs public transportation users. The analysis of effects on traffic density, fuel consumption and travel time help to explain the variables and their relationships to achieve a sustainable transport system. The variable space, the trade-off from different perspectives and geo-location analysis give a comprehension to the decision makers to gain knowledge about the model.

MOEA Adaptive ε-Sampling and v-Hood (AεSvH): The AεSvH algorithm, an extension of the multi- and many-objective AεSεH algorithm. The proposed AεSvH creates neighborhoods in variable space to bias mating for recombination and uses ε-dominance principles to truncate the population when the number of non-dominated solutions is larger than the population size. The difference between the proposed AεSvH algorithm and the AεSεH is the space where neighborhoods for parent selection is performed. In AεSεH neighborhoods are created in objective space, whereas in AεSvH neighborhods are created in decision space. Specifically, we use AεSvH in bi-objective optimization problem of Level of Service (LoS). We use variable nPt which defines the proportion of public transportation users, to create the neighborhoods so that all solutions with the same proportion of public transportation users belong to the same neighborhood. In this way, the method allows a less disruptive recombination compared to AεSεH for the problem dealt with in this work.

1.4 Outline

Chapter 2 describes the optimization framework and its components. A description of each model's component is given according to the next order: Problem formulation, Simulation, Evolutionary Algorithms, Analysis of Solutions and Design Knowledge. Problem formulation refers to the implementation model. Simulation refers to the traffic simulator used to test the

mobility scenarios. Evolutionary Algorithms (EA) introduces the main concepts, definitions and EAs implemented. Analysis of Solutions are the methods that we use to study the optimized solutions, the variable, objective and geographical space. Finally, a description of Design of Knowledge is given.

Chapter 3 describes the area of the study and the main components of the mobility and transport scenario: the network infrastructure and the mobility model based on activities.

Chapter 4 is devoted to the traffic signal optimization problem since a single objective approach. It includes the problem formulation, the EA algorithm, the problem representation and the operators designed and implemented. The optimization problem is addressed considering the travel time as a fitness function. A complete and comprehensive analysis of results is performed and discussed. Finally, conclusion and future work are presented.

Chapter 5 describes a bi-objective optimization approach for traffic signals considering travel time and fuel consumption as objective functions. The trade-off is analyzed and discussed in the results section.

Two additional chapters examine a different problem regarding Level of Service optimization based on traffic density levels in urban transportation. A multi-objective evolutionary algorithm searches the combinations of proportion private/public transport users, capacity and headways of Bus Rapid Transit (BRT) of five main corridors in the area of study. Different proportions and configurations impact the traffic density, travel time and fuel consumption.

In chapter 6 a bi-objective approach studies the trade-off between traffic density and travel time. The method, the evolutionary algorithm, the representation and operators are described. A complete trade-off analysis is conducted in the respective chapter sections. Two BRT optimized configurations are studied to analyze the effects of fuel consumption and emissions in the geographical area.

Chapter 7 study the Level of Service optimization from a three-objective approach considering travel time, traffic density and fuel consumption. The chapter describes the method, the evolutionary algorithm, the operators, the results and discussion.

Finally, Chapter 8 summarizes this work, presents the conclusions and propose a future research.

Chapter 2

Design Optimization Framework

This chapter presents the framework and its components: Problem formulation, Simulation, Evolutionary Algorithms, Analysis of Solutions and Design Knowledge. The chapter includes a description of each component, the simulators, the evolutionary algorithms and definitions.

2.1 Introduction

The central purpose of this work is to implement a design optimization framework which applies the evolving design concept. Evolving designs on computers (EDC) empower us to employ computer models at every stage of the design process improving the quality, productivity, speed and decreasing the expense of design. EDC is based on evolutionary biology principles to generate elaborate and innovative designs (Bentley, 1999). The proposed design optimization framework is based on evolutionary design optimization, which starts with an existing design and evolves those variables or parameters of the design which we are searching for improvement. Starting from a problem definition in terms of search-space, an evolutionary algorithm (EA) uses an analogy with natural evolution to perform the search by evolving solutions to problems. The variables or parameters are encoded in a specific problem-representation by a genetic code which is evolved by an EA. The EA guide the evolution towards better areas of the search space by evaluating every solution in the population by a fitness function which assesses how well the solution satisfies the problem. Those solutions that represent promising designs are analyzed and judged by the experts who accept the proposed design or suggest a new set of design parameters to be evolved.

2.2 Components

We follow the design optimization approach illustrated in Fig 2.1. We first formulate a problem related to mobility and transport, simulate the movements in the city using a specialized transport simulator. We use a single and multi-objective evolutionary algorithm to find optimal solutions to the formulated problem evaluating the quality of solutions using the outcome of the simulation, analyze the solutions produced by the evolutionary algorithm and extract knowledge about the system. The approach is iterative,

where the solutions and obtained knowledge are made available to the expert for further analysis. The feedback from the expert could be used to reformulate the problem to study additional details. Once the expert is satisfied, the knowledge gained could be suggested to improve the real system.

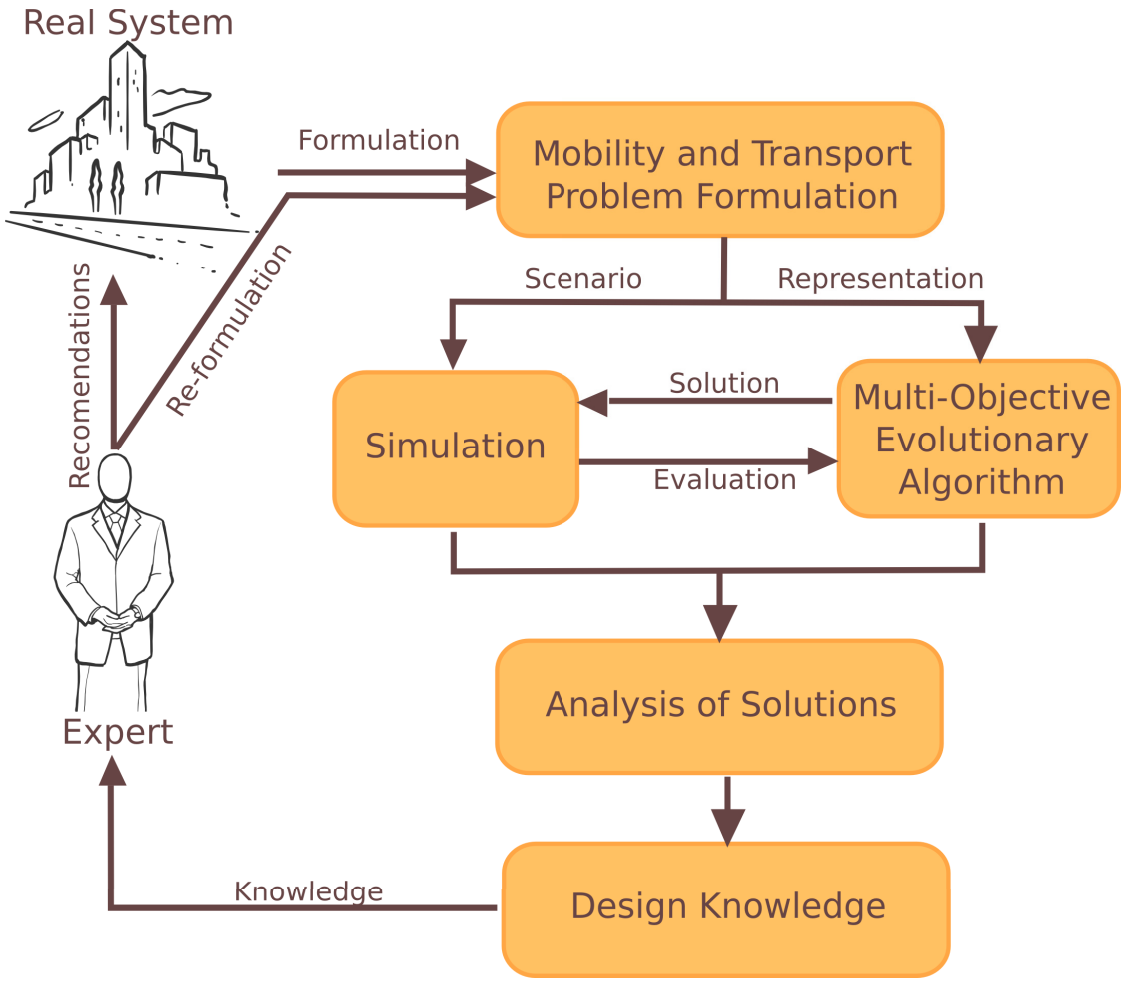


FIGURE 2.1: Method

2.3 Problem Formulation

The problem formulation is an abstraction or simplification of the real-world problem, and it represents the model. The real-world solving problem could be a challenging task, especially when those problems remain complicated due to the coincidence of several factors. One factor is that the search space could be broad that the exhaustive solution exploration is denied. Another one refers to the evaluation function; it could be hard to find a complete formulation. Moreover, some hard constraints can restrict feasible solutions search. In this work, we use models that consider those factors, by using evolutionary algorithms (EA) to explore the search space and simulators to evaluate the solution.

2.4 Simulation

Computer simulation refers to the use of computational models to gain additional insight into a complex system's behaviour. Furthermore, those computational models allow the evaluation of designs and plans without bringing them to the real world (e.g. road networks and traffic lights). Synthetic environments are in most cases a necessity since the simulated system is non-observable (currently being designed), unsafe or unethical (e.g. humans or animals involved could be harm) or directly unpractical (experiment or data acquisition costs, change in the systems observed is slow). Simulation is not a decision-making tool but a decision support tool, allowing better-informed decisions to be made. For this work, we use mobility and traffic simulators to implement the mobility model and emission simulator to compute the fuel consumption and gas emissions. Next section describes each of them.

2.4.1 Multi-Agent Transport Simulation MATSim

Matsim comes from Multi-Agent Transport Simulation, provides a framework to implement large-scale agent-based transport simulations(Horni, Nagel, and Axhausen, 2016). It uses a microsimulation approach. The term microsimulation relates to the microscopic (individual) demand of each person in the scenario. MATSim works in four main phases. In the first, a synthetic population is generated, which is then used to generate the initial travel demand (initial mobility plans). The travel demand is then optimized, and finally, results are analyzed (see Fig.2.2). Next section describes the Matsim process(Horni, Nagel, and Axhausen, 2016).

- Scenario Creation: It refers to set the geographical area and geographical data sources, and other information like census data or mobility studies that can help to model the mobility scenario. An important topic is the transportation network modelling, for that purpose, open geographic data sources have been used.

- Initial Individual Demand Modeling: Here, set the initial mobility plan for each agent. A mobility plan defines activities that each person will perform. Major activity attributes are type, location, duration, opening and closing times.

- Demand Execution: In this process, the simulator uses all plans for all individuals to simulate the traffic during 24 hours. The simulation is based on a queue model and uses a time step based approach with a step width of one second, meaning that the system state is calculated every second.

- Scoring: To improve the mobility plan in each iteration, it uses a scoring function. The score function rate the quality of the day plans in the population. The scoring function has two main components: the first part evaluates accomplished activities, a positive contribution for the (usually) positive utility earned by performing an activity. The function

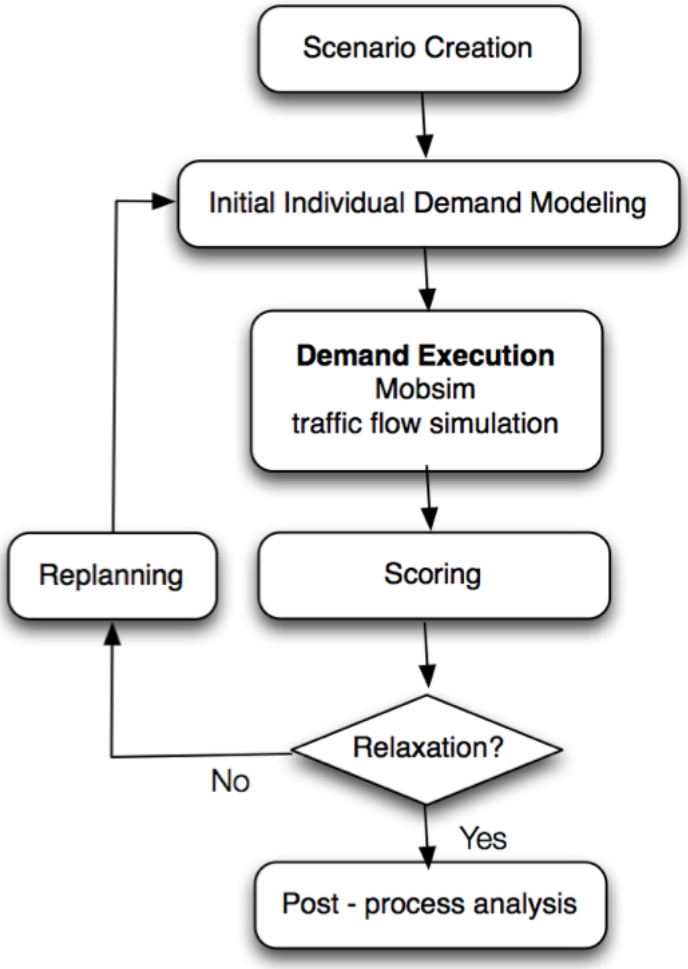


FIGURE 2.2: Matsim process

for activity evaluates duration for each activity and penalize for waiting time, late arrival, early departure and short duration. The second part evaluates the movements. It is a negative contribution (penalty) for travelling.

- Replanning: Each re-planning module takes charge for a specific part in the optimization process. As an example, the *Router* module calculates the routes of a plan based on the amount of traffic from the last traffic flow simulation. The *Time Allocation Mutator* module modifies departure times and activity durations of a daily plan. This module varies the corresponding times randomly. Additional modules could change activities' locations, or change the sequence of activities. A characterization of modules is whether they modify a plan randomly (Random Mutation) or whether they search for the best solution based on the results of the last traffic flow simulation (Best Response).

- Relaxation: The relaxed state of the co-evolutionary algorithm of MATSim is reached if the utility for each agent does not noticeably change

through variation of the day plans. The simulator evaluates the average worst plan for all agents and compares with the average best plan if there is not a considerable difference, the iteration ends. This condition is known as *relaxation* or *equilibrium state*. The relaxed state of the co-evolutionary algorithm of MATSim is reached if the utility for each agent does not noticeably change through variation of the day plans.

- Post process analysis: The simulator provides as output a detailed log of the movement and time of all agents. This log is used for analyzing the mobility and the status of the infrastructure around the city.

2.4.2 Modules

Emissions

MATSim emission extension (Hülsmann et al., 2011) computes the gas emissions per link per agent, at the time an agent enters a link of the transport network. It calculates warm and cold-start exhaust emissions for cars by connecting MATSim simulation output to the detailed database Handbook on Emission Factors for Road Transport (HBEFA) (Keller and Wuthrich, 2014).

MATSim computes warm emissions deriving the kinematic characteristics from the simulation and combines this information with vehicle characteristics to extract emission factors from the database of the Handbook on Emission Factors for Road Transport (HBEFA). To derive the kinematic characteristics, the emission model considers 'free flow' and 'stop&go' as traffic states per road segment. To calculate cold-start emissions, MATSim derives parking duration and accumulated traveled distance from the simulation. For parking duration, HBEFA database differentiates emission factors in one hour time steps from 1h to 12h. After 12 hours the vehicle is fully cooled down. There are also different cold emission factors for short trips (less than 1Km) and longer trips (greater than 1Km)(Kickhöfer et al., 2013).

Traffic Signals

MATSim uses an specific configuration to simulate traffic lights microscopically using fixed-time control. MATSim uses an structure displayed in the figure 2.3, given by *Signal System, Signal Group, Control and Signal*.

- Signal System is a collection of signal groups that is controlled by the same group.
- Signal Group refers to a logical group of the traffic lights, all lights display same color at the same time. A group as a set of lights that control a specific flow or movements: for example North-South or West-East.
- Control: Algorithm or control scheme that determines which colors are displayed by the different signal groups.

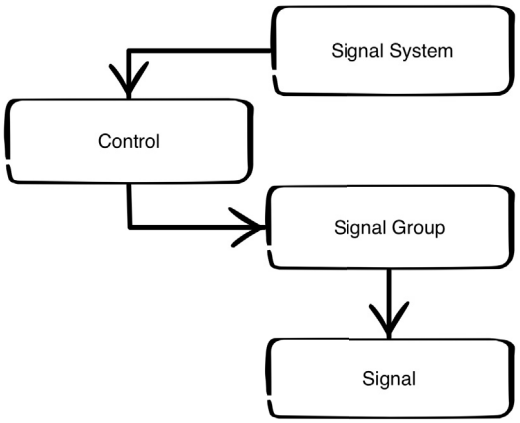


FIGURE 2.3: Traffic Light MATSim Representation

- Signal is the traffic light or traffic signal located at specific place in the network.

2.4.3 Comprehensive Modal Emission Model (CMEM)

CMEM is a microscopic emissions simulator that computes second-by-second tailpipe emissions and fuel consumption based on different vehicle operating modes (modal), such as idle, steady-state, cruise, and various levels of acceleration/deceleration (G.Scora and Barth, 2006). It is called comprehensive because it can predict emissions for a wide range of vehicle / technology categories and various operating conditions, such as properly functioning, deteriorated, malfunctioning. CMEM requires two groups of inputs, input operating variables and model parameters. The input operating variables include information about the activity of the vehicles, that is second-by-second speed (from which acceleration can be derived) and the slope of the road. The model parameters are determined for each one of the vehicles used in the simulation according to the categorization established by CMEM.

2.5 Evolutionary Algorithms

2.5.1 Main Concepts and Definitions

Evolutionary Algorithms (EA) are a set of *meta-heuristic* problem-solvers which guide to lower-level heuristics to search feasible solutions in complex search land-scapes (Coello, Lamont, and Veldhuizen, 2006). A heuristic is a technique in which comparison rules select the most appropriate partial or local solution. EA comprises genetic algorithms (GA), evolution strategies (ES), and evolutionary programming (EP). In this work, we focus in single and multi-objective EAs as a meta-heuristic model to search the feasible solutions of the real-world problem that we model.

Because of the complexity of the problem owing to either the number of possible solutions in the *search space* is vast or *evaluation function* that describes the quality of any proposed solution is noisy or varies with time or has constraints that require special operations(Michalewicz and Fogel, 2010), EA algorithms are a set of tools that facilitate the exploration of vast spaces, not just a single solution but an entire series of solutions.
EA techniques are based on the Darwinian evolution concept which relies on natural selection and the principle of survival of the fittest. It consists of a population (of encoded solutions) shaped by a set of mutation and crossover operators and evaluated by a fitness function that determines which survive into the next generation.

Algorithm 1: Evolutionary Algorithm

1: $P \leftarrow$ Initial Population()
2: Evaluation (P)
3: **repeat**
4: $P\prime \leftarrow$ Parent selection(P)
5: $Q \leftarrow$ Recombination and mutation ($P\prime$)
6: Evaluation (Q)
7: $P \leftarrow$ Survival selection (P,Q)
8: **until** condition is met

A simple EA is shown in Alg.1. A population P of individuals is initialized, and then each is evaluated by a fitness function $Evaluation(P)$ according to its worth in some domain or problem. Fitness evaluation can be a simple formulation or a more complex evaluation by an elaborate simulation. After that, an iterative process where parent selection, recombination and mutation, evaluation and survival selection is performed generation by generation until a specific condition such as a number of generations or evaluations is met. Parent selection decides which one becomes a parent and the number of children that they can breed. Individuals with higher fitness are more likely to be parents. Recombination and mutation operators create offspring Q. Recombination allows interchanging the genetic code between parents and mutation alter the offspring. The offspring are then evaluated. Finally, the survival step decides which one persists in the population. A variation in this step is to consider *elistism* that allows to the best individual always to survive, ensuring that when optimum is found it cannot be lost.

2.5.2 Single Objective Optimization

Definition 1 (Single-Objective Optimization Problem) *Given a vector* $\vec{x} = [x_1, \ldots, x_n]^T$ *, of n decision variables* x_i*, that fulfills the following* $r = a + b$ *restrictions:*

$$\vec{g}(\vec{x}) = [g_1(\vec{x}), ..., g_a] \geq 0 \tag{2.1}$$

$$\vec{h}(\vec{x}) = [h_1(\vec{x}), ..., h_b] = 0 \tag{2.2}$$

and minimize (or maximize) the function:

$$\vec{f}(\vec{x}) \tag{2.3}$$

Definition 2 (Single-Objective Global Minimum Optimization) *Given a function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}, \Omega \neq \varnothing$, for $x \in \Omega$ the value $f^* \triangleq f(x^*) > -\infty$ is called a global minimum if and only if:*

$$\forall x \in \Omega : f(x^*) \leq f(x) \tag{2.4}$$

x^* is by definition the global minimum solution, f is the objective function, and the set Ω is the feasible region of x. The goal of determining the global minimum solution(s) is called the global optimization problem for a single objective problem.

2.5.3 Multi-Many Objective Optimization

In particular, multi-many objective EA (MOEA) possess characteristics that are beneficial for problems involving i) multiple conflicting objectives, and ii) intractably large and highly complex search spaces(Eckart, 2004). In this work we use the Adaptive ε-Sampling and ε-Hood (AεSεH) (Aguirre, Oyama, and K., 2013) algorithm to search optimal solutions. AεSεH is an elitist evolutionary multi- and many-objective optimizer that applies ε-dominance principles both for survival selection and parent selection. In the following, we describe the main features of the algorithm, representation, operators of variation, and fitness functions used to study our system.

Following section present generic multi-objective problem (MOP) and Pareto optimally mathematical and formal symbolic definition

Definition 3 (Multi-Objective Optimization Problem) *Given a vector $\vec{x} = [x_1, \ldots, x_n]^T$, of n decision variables x_i, that fulfills the following $r = a + b$ restrictions:*

$$\vec{g}(\vec{x}) = [g_1(\vec{x}), ..., g_a] \geq 0 \tag{2.5}$$

$$\vec{h}(\vec{x}) = [h_1(\vec{x}), ..., h_b] = 0 \tag{2.6}$$

and optimizes the vector of functions:

$$\vec{f}(\vec{x}) = [f_1(\vec{x}), ..., f_m]^T \tag{2.7}$$

Optimize here could indicate minimizing or maximizing the values of each objective functions which is subject to the problem at hand. A Multi-Objective Optimization Problem becomes a Many-Objective *one when, according to the above definition, $m \geq 4$ (Coello, Lamont, and Veldhuizen, 2006).*

The concept of Pareto optimum was formulated by Vilfrido Pareto during the XIX century and marks the origin of the research on Multi-Objective Optimization (Coello, Lamont, and Veldhuizen, 2006). In this section will be defined the concepts of Pareto dominance, Pareto optimality, Pareto set and Pareto front (**Veldhuizen99**; **ToscanoPulido01**).

Definition 4 (Pareto dominance) *Given two decision vectors* $\vec{u} = (u_1, \ldots, u_n)$, $\vec{v} = (v_1, \ldots, v_n)$*:*

$$\begin{array}{ll} \vec{u} \prec \vec{v} & \text{if and only if } f_i(\vec{u}) < f_i(\vec{v}) \\ \vec{u} \text{ dominates } \vec{v} & \forall i \in (1, \ldots, k) \end{array}$$

$$\begin{array}{ll} \vec{u} \preceq \vec{v} & \text{if and only if } f_i(\vec{u}) \leq f_i(\vec{v}) \\ \vec{u} \text{ weakly dominates } \vec{v} & \forall i \in (1, \ldots, k) \end{array}$$

$$\begin{array}{ll} \vec{u} \sim \vec{v} & \text{if and only if } f_i(\vec{u}) \not\leq f_i(\vec{v}) \wedge f_i(\vec{u}) \not\leq f_i(\vec{v}) \\ \vec{u} \text{ is not comparable to } \vec{v} & \forall i \in (1, \ldots, k) \end{array}$$

The previous operators are given in function of a minimization problem, nevertheless the definitions are similar to the ones for maximization problems ($\succ, \succeq, \sim$).

Definition 5 (Pareto optimality) *Given a decision vector* $\vec{u}$*, it can be said that its Pareto optimal if and only if:* $\vec{u} \in \Omega \mid \neg\exists\ \vec{v} \in \Omega \mid \vec{f}(\vec{v}) \preceq \vec{f}(\vec{u})$

A decision vector that its Pareto optimal cannot be improved in any of the objectives without this inflicting a degradation in at least one other objective.

Definition 6 (Pareto optimal set) *Is defined as* $P = \{\vec{u} \in \Omega \mid \neg\exists\ \vec{v} \in \Omega \mid \vec{f}(\vec{v}) \preceq \vec{f}(\vec{u})\}$.

The Pareto optimal set is part of the decision variable space

Definition 7 (Pareto front) *Given a Multi-Objective Optimization problem* $\vec{y} = \vec{f}(\vec{x})$ *and a Pareto optimal set* P^**, the Pareto front* (PF) *its defined as:*

$PF = \{\vec{y} = \vec{f} = (\vec{f}_1(\vec{x}), ..., \vec{f}_k(\vec{x}) \mid \vec{x} \in P\}$

The Pareto front is part of the solutions space.

2.5.4 EA used in this work

Elitism Single Objective EA

We use a single-objective EA for the traffic signal optimization problem. Traffic signal functioning is defined by several parameters. First, the cycle length. Then, the green times of the different phases for the intersecting lanes. Finally, the offset between the beginning of the cycles of consecutive signals to coordinate them and induce continuous flows. Therefore, the search space for the optimal setting for a set of traffic lights grows rapidly with the number of signals being considered. Evolutionary algorithms (EAs) provide the means to search and explore several alternatives, allowing the human expert to learn about the solution space and focus on the most suitable solutions.
Fig. 2.4 illustrates the flowchart of the EA algorithm. In the following, we describe the main steps of the algorithm based on this figure. Each solution is a traffic signal coded by its representation.

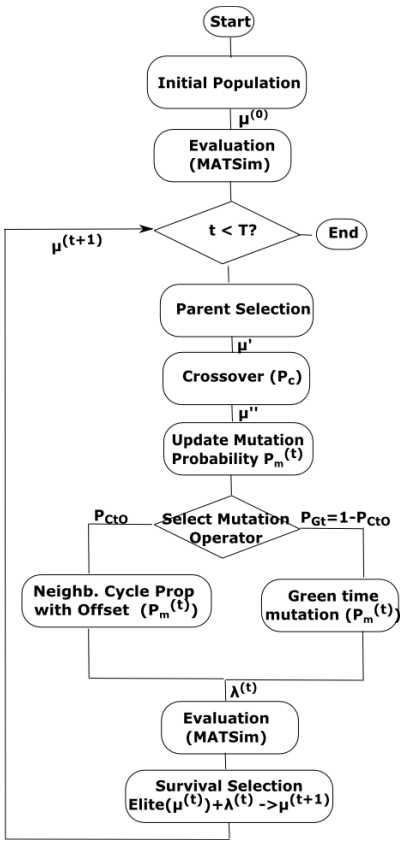


FIGURE 2.4: Evolutionary Algorithm Flow Chart

The initial population μ is formed by 20 deterministically created solutions. No repetitions are allowed. We do this to guarantee maximum diversity for each signal across the population so that, evolution could explore a broad region of variable space. Each solution is evaluated running MATSim with the traffic signals settings it encodes. Parents are selected for reproduction using binary tournaments among randomly sampled solutions. The offspring population λ is created applying crossover to the selected parents with probability P_c followed by mutation. Mutation updates the probability $P_m^{(t)}$, selects a mutation operator and apply it with the updated $P_m^{(t)}$ per signal. There are two mutation operators, one for cycle time and offset and another for green time. The operators' probabilities are P_{CtO} and $P_{Gt} = 1 - P_{CtO}$, respectively. During the evolution we relax the initial assumption of common cycle length for all signals in a solution. The proposed mutation operators for cycle time and offset searches for groups of signals with same cycle length by propagating the cycle of a signal to its neighbors and induce their coordination by setting their offset based on the distance between neighboring signals. The operators' probabilities allows to balance the search for cycle lengths and green times.

After crossover and mutation, the offspring is evaluated running MATSim as indicated above. Selection of solutions for the next generation applies elitism, taking into consideration the current population $\mu^{(t)}$ and the offspring population $\lambda^{(t)}$. Namely, the population $\mu^{(t+1)}$ is formed by the

best in the combined population of the top *nElite* from μ_t and the λ_t offspring, i.e. $nElite(\mu^{(t)}) + \lambda^{(t)} \rightarrow \mu^{(t+1)}$. Also, we introduce varying mutation rates for $P_m^{(t)}$. The algorithm evolves solutions repeating parent selection, crossover and mutation, offspring evaluation, and survival selection until a pre-determined number of generations T has been reached. The appropriate combination of operators, high selection pressure via elitism, and varying mutation rates accelerate convergence of the algorithm and allows to search effectively even with small populations.

MOEA:Adaptive ε-Sampling and ε-Hood AεSεH Algorithm

AεSεH follows the main steps of a population-based evolutionary algorithm, i.e. parent selection, offspring creation and survival selection, adjusting its operation depending on whether the population contains dominated solutions or not.

To perform survival selection, the current population and its offspring are combined and divided into non-dominated fronts using the non-dominated sorting procedure. If the number of non-dominated solutions in the first front is smaller than the population size, the sorted fronts of non-dominated solutions are copied one at the time to the next population until it is filled; if the last copied front overfills the population, the required number of solutions are chosen randomly from it to have the exact number specified by the population size. On the other hand, if the number of non-dominated solutions in the first front is larger than the population size, the first front is truncated to the size of the population using the *ε-sampling* procedure. *ε-sampling* randomly chooses solutions from the first front to include them in the surviving population, eliminating from the front those solutions that are ε-dominated by the chosen samples. As a result, solutions in the next population are spaced according to the $\vec{f}(\vec{x}) \mapsto^{\varepsilon_s} \vec{f}'(\vec{x})$ mapping function and parameter ε_s used to compute ε-dominance between solutions.

For parent selection, the algorithm first uses a procedure called *ε-hood creation* to cluster solutions in objective space and then applies *ε-hood mating* to select parents. When all solutions in the population are non-dominated, *ε-hood creation* selects *randomly* an individual from the population and applies ε-dominance with mapping function $\vec{f}(\vec{x}) \mapsto^{\varepsilon_h} \vec{f}'(\vec{x})$ and parameter ε_h. A neighborhood is formed by the selected solution and its ε_h-dominated solutions. Neighborhood creation is repeated until all solutions in the population have been assigned to a neighborhood. *ε-hood mating* sees the neighborhoods as elements of a list and visits them one at the time in a round-robin schedule. The first two parents are selected *randomly* from the first visited neighborhood in the list. The next two parents are selected randomly from the second neighborhood in the list, and so on. When the end of the list is reached, parent selection continues with the first neighborhood in the list. On the other hand, when dominated solutions are present in the population, *ε-hood creation* makes sure that the solution sampled to create the neighborhood is a non-dominated solution and *ε-hood mating* uses *binary tournaments* based on dominance rank to select parents within the neighborhoods. Both epsilon

Algorithm 2: AεSvH

Require: Population size P_{size}
Ensure: $\mathcal{F}_1$, set of Pareto non-dominated solutions

1: $\varepsilon_s \leftarrow \mathbf{0}$, $\Delta_s \leftarrow \Delta_\mathbf{0}$ {set ε_s-dominance factor and its step of adaptation}
2: $\mathcal{P} \leftarrow random$ {initial population $\mathcal{P}$, $|\mathcal{P}| = P_{size}$}
3: evaluation($\mathcal{P}$)
4: **repeat**
5: {Parent selection}
6: $\{\mathcal{H}, N_H\} \leftarrow v$-hood creation ($\mathcal{P}$) $\{\mathcal{H} = \{\mathcal{H}_j\},\ j = 1, 2, \cdots, N_H\}$
7: $\mathcal{P}' \leftarrow v$-hood mating($\mathcal{H}$, $Psize$)
8: {Offspring creation}
9: $\mathcal{Q} \leftarrow$ recombination and mutation($\mathcal{P}'$) $\{|\mathcal{Q}| = |\mathcal{P}| = P_{size}\}$
10: {Evaluation and front sorting}
11: evaluation($\mathcal{Q}$)
12: $\mathcal{F} \leftarrow$ non-dominated sorting($\mathcal{P} \cup \mathcal{Q}$)
$\{\mathcal{F} = \{\mathcal{F}_i\},\ i = 1, 2, \cdots, N_F\}$
13: {Survival selection}
14: $\{\mathcal{P}, N_S\} \leftarrow \epsilon$-sampling truncation($\mathcal{F}$, ϵ_s, P_{size}) {N_S, number of samples}
15: $\{\epsilon_s, \Delta_s\} \leftarrow$ adapt (ϵ_s, Δ_s, P_{size}, N_S)
16: **until** termination criterion is met
17: **return** $\mathcal{F}_1$

parameters ε_s and ε_h used in survival selection and neighborhood creation, respectively, are dynamically adapted during the run of the algorithm. This algorithm has been shown to work effectively on continuous and discrete multi- and many-objective optimization problems (Aguirre, Oyama, and K., 2013) (Aguirre et al., 2014b) (Aguirre et al., 2014a). Further details about the algorithm can be found in (Aguirre, Oyama, and K., 2013) and (Aguirre et al., 2014b).

MOEA: Adaptive ε-Sampling and v-Hood (AεSvH) Algorithm

In this section we explain the proposed AεSvH algorithm, an extension of the multi- and many-objective (AεSεH) algorithm (**Aguirre2014**; Aguirre, Oyama, and K., 2013). The proposed AεSvH creates neighborhoods in variable space to bias mating for recombination and uses ε-dominance principles to truncate the population when the number of non-dominated solutions is larger than the population size.

The algorithm's flow is illustrated in **Procedure 2**. It starts by setting initial default values for the parameter ε_s used in and its step of adaptation Δ_s. The population $\mathcal{P}$ is randomly initialized and evaluated. The main loop starts by creating neighborhoods of solutions, grouping them by proximity in

variable space using a v-hood creation procedure. Next, v-hood mating creates a pool of mates $\mathcal{P}'$ selecting them from the same neighborhood. The selected mates are recombined and mutated to create the offspring population $\mathcal{Q}$. Since mates come from the same neigborhood, a solution would recombine only with a solution that is close by in decision space. After offspring $\mathcal{Q}$ is evaluated, non-dominated sorting is performed on the population that results from joining the current population $\mathcal{P}$ and its offspring $\mathcal{Q}$. The population of size $2P_{size}$ sorted in non-dominated fronts $\mathcal{F}$ is then truncated to obtain the surviving population $\mathcal{P}$ of size P_{size} using a ε-sampling truncation procedure set with parameter ε_s. The parameter ε_s and its step of adaptation Δ_s are adapted every generation so that the number of ε-sampled solutions approach the size of the population. In the following we explain with more detail v-hood creation, v-hood mating and ε-sampling truncation.

The v-hood creation procedure splits the surviving population $\mathcal{P}$ into neighborhoods as illustrated in **Procedure 3**. First, this procedure randomly selects an individual z from the population $\mathcal{P}$. Then, solutions in $\mathcal{P}$ with the same proportion of public transportation users as $z_{N_{Pt}}$ are grouped in $\mathcal{Y}$. Next, z and $\mathcal{Y}$ are assigned to neighborhood $\mathcal{H}_i$ and removed from the population $\mathcal{P}$. Neighborhood creation is repeated until all solutions in $\mathcal{P}$ have been assigned to a neighborhood.

Mating for recombination is implemented by the procedure v-hood mating illustrated in **Procedure 4**. Neighborhoods are considered to be elements of a list. To select two mates, first a neighborhood from the list is specified deterministically in a round-robin schedule. Then, two individuals are selected randomly within the specified neighborhood, so that an individual will recombine with other individual that is located close by in variable n_{Pt} space. Due to the round-robin schedule, the next two mates will be selected from the next neighborhood in the list. When the end of the neighborhood lists is reached, mating continues with the first neighborhood in the list. Thus, all individuals have the same probability of being selected within a specified neighborhood, but due to the round-robin scheduling individuals belonging to neighborhoods with fewer members have more recombination opportunities that those belonging to neighborhoods with more members. Once the pool of all mates $\mathcal{P}'$ has been established, they are recombined and mutated according to the order they were selected during mating.

The ε-sampling truncation (Aguirre, Oyama, and K., 2013) receives the sets of solutions $\mathcal{F}$ created by non-dominated sorting and selects exactly P_{size} surviving solutions from them. If the number of non-dominated solutions $|\mathcal{F}_1| < P_{size}$, the fronts $\mathcal{F}_i$ are copied iteratively to $\mathcal{P}$ until the population is filled. If $\mathcal{F}_i$ overfills $\mathcal{P}$, the required number of solutions are selected randomly from $\mathcal{F}_i$ and copied to $\mathcal{P}$. On the other hand, if the number of non-dominated solutions $|\mathcal{F}_1| > P_{size}$, it calls ε-sampling with parameter ε_s to get from $\mathcal{F}_1$ a sample of ε_s non-dominated solutions $\mathcal{F}_1^{\varepsilon} \leq \mathcal{F}_1$ and the set of ε_s-dominated solutions $\mathcal{D}^{\varepsilon_s}$. $\mathcal{F}_1^{\varepsilon}$ is copied to $\mathcal{P}$. If $|\mathcal{P}| > P_{size}$, solutions are randomly eliminated from $\mathcal{P}$ until its size is P_{size}. If $|\mathcal{P}| < P_{size}$, solutions are randomly copied from the ε_s-dominated solutions $\mathcal{D}^{\varepsilon_s}$ to $\mathcal{P}$ until its size is P_{size}. The ε sampling procedure is done in way that the surviving

population $\mathcal{P}$ always includes the extreme solutions present in $\mathcal{F}_1$.

Algorithm 3: v-hood creation ($\mathcal{P}$)

Require: Population $\mathcal{P}$
Ensure: Neighborhoods $\mathcal{H} = \{\mathcal{H}_i\}, i = 1, 2, \cdots, N_H$

$\mathcal{H} \leftarrow \emptyset$
$i \leftarrow 0$
while $\mathcal{P} \neq \emptyset$ **do**
  $z \leftarrow x_r \in \mathcal{P} \mid r = rand(\ 1,\ \mathcal{P}\ )$ {z, is a randomly chosen solution from list}
  $\mathcal{Y} \leftarrow \{y \in \mathcal{P} \mid z_{N_{Pt}} == y_{N_{Pt}}\}$ { solutions which N_{Pt} variable are equal }
  $i \leftarrow i + 1$
  $\mathcal{H}_i \leftarrow \{\{z\} \cup \mathcal{Y}\}$ { z and its solutions form the N_{Pt}-hood }
  $\mathcal{H} \leftarrow \mathcal{H} \cup \mathcal{H}_i$
  $\mathcal{P} \leftarrow \mathcal{P} \setminus \mathcal{H}_i$
end while
$N_H \leftarrow i$
return $\mathcal{H}$, N_H

Algorithm 4: v-hood mating ($\mathcal{H}$, P_{size})

Require: Neighborhoods $\mathcal{H} = \{\mathcal{H}_i\}, i = 1, 2, \cdots, N_H$, and population size P_{size}
Ensure: Pool of mated parents $\mathcal{P}'$, $|\mathcal{P}'| = 2P_{size}$

$\mathcal{P}' \leftarrow \emptyset$
$i \leftarrow 1$
$j \leftarrow 0$
while $j < P_{size}$ **do**
  $\{x_{r_1}, x_{r_2} \in \mathcal{H}_i \mid r_1 \wedge r_2 = rand(\ 1,\ |\mathcal{H}_i|\ ), r_1 \neq r_2\}$
  $y \leftarrow tournament(x_{r_1}, x_{r_2})$ {d}ecide based on dominance rank
  $\{x_{r_3}, x_{r_4} \in \mathcal{H}_i \mid r_3 \wedge r_4 = rand(\ 1,\ |\mathcal{H}_i|\ ), r_3 \neq r_4\}$
  $z \leftarrow tournament(x_{r_3}, x_{r_4})$ {d}ecide based on dominance rank
  $\mathcal{P}' \leftarrow \mathcal{P}' \cup \{y, z\}$
  $i \leftarrow 1 + (i \bmod N_H)$
  $j \leftarrow j + 1$
end while
return $\mathcal{P}'$

The difference between the proposed AεSvH algorithm and the AεSεH is the space where neighborhoods for parent selection is performed. In A$\boldsymbol{\varepsilon S\varepsilon H}$ neighborhoods are created in objective space, whereas in A$\boldsymbol{\varepsilon SvH}$ neighborhods are created in decision space. Specifically, we use the variable $\boldsymbol{n_{Pt}}$ to create the neighborhoods so that all solutions with the same proportion of public transportation users belong to the same neighborhood. In this way the method allows a less disruptive recombination compared to A$\boldsymbol{\varepsilon S\varepsilon H}$ for the problem dealt with in this work.

2.6 Analysis of Solutions

Qualitative data analysis is a process that pretends to simplify and make sense of a massive amount of information from diverse sources. It allows to take descriptive information and offer an explanation or interpretation. The proposed framework generates several outcomes generated by simulation and MOEA. Agent-based transport simulation models integrate and manage several data sources (e.g. spatial data, activity diary trips, agent attributes). EA (single- and multi-objective) gives the set of solutions (variables and objectives) during the evolution. In this work, the analysis focus in variable space, objective space and spatial geo-location space. For variable and objective space, the study is centred in the range of the variables and objectives. A complementary analysis is conducted to seek an additional meaning, explanation and interpretation in the context of the problem by a geo-location of the evolved solutions into the geographical area of study.

2.7 Design Knowledge

Our framework aims to be a support and add value to the body of knowledge during the design process for different complex problems. For this work, the central intention of the evolving design is to provide the designer with a set of potential solutions so that he/she has a plethora of candidates solutions from the optimal set. Moreover, show a path to go from the current to a potential optimized state. In many-objective optimization problems with trade-offs, the framework reveals the optimal solution set, so that, the decision maker decides to choose acceptable solutions based on his/her knowledge and expertise.

Finally, considering additional criteria from the expert, the proposed problem modelling and its solutions will be evaluated and verified so that a new version of the problem can be implemented to accomplish a new cycle iteratively.

Chapter 3

Quito Model Mobility Scenario

This chapter describes the area of study used for this research. A brief introduction about Quito city and its characterization and the mobility scenario: the transport network and mobility plans.

3.1 Quito Metropolitan District-DMQ

Quito Metropolitan District (DMQ) is the capital city of Ecuador and located in the central part of the Ecuador. Located at 2850 meters above sea level in the Andes mountains is the second one highest cities around the world. Fig. 3.1 display a view of the DMQ urban extension.

FIGURE 3.1: DMQ Urban Area (captured by google earth)

DMQ's urban population growth has put enormous strains on its transportation system. There is the need to model the system, better understand the problems associated with transportation and mobility, find sustainable solutions to solve them and improve life in the city. DMQ's urban population increased from 1 million in 1990 to 1.7 million in 2015, and it is expected to increase other 0.5 million by 2030 (United Nations and Social Affairs, 2014).

TABLE 3.1: Reasons for trips in DMQ

Activity	% Population
Study	32.5
Work	31.1
Personal business	24.3
Shopping	4.4
Medical Visits	3.1
Leisure/Sports	1.6
Others	2.9

DMQ concentrates several activities: economic, industrial, services, administrative and political, due to of its importance as one of the Ecuador's main cities.
DMQ keep a constant and intense interchange traffic flows inside and outside of its metropolitan area. Outside of the metropolitan area, with other cities due to activities such as business and tourism. Inside of the metropolitan area, people continuously move mainly to accomplish several tasks. According to a mobility survey (DMQ, 2012) the main reasons for trips in DMQ are study, work, personal business, leisure, sports and others. Table 3.1 show the population percentage for each activity category. DMQ's urban area has a high concentration of population, around 77% of the total population live in the urban area.

In the DMQ the 84.4% of the trips are by using cars or bus. From this percentage, 74% are in public transportation and the rest in private cars (DMQ, 2012). In Ecuador, the rate of vehicles for each 1000 inhabitants changed from 17 to 51 from 1990 to 2010 (INEC, 2010a). Fig. 3.2 shows the inhabitants against the number of vehicles in Ecuador during 40 years since 1968. Note a distinct tendency of increase in the number of vehicles according to the population growth. Inside of DMQ, the average rate of vehicles per family is 0,51 that represents approximately 0.13 vehicles per inhabitant and more than 90% use their homes as parking places(DMQ, 2012).

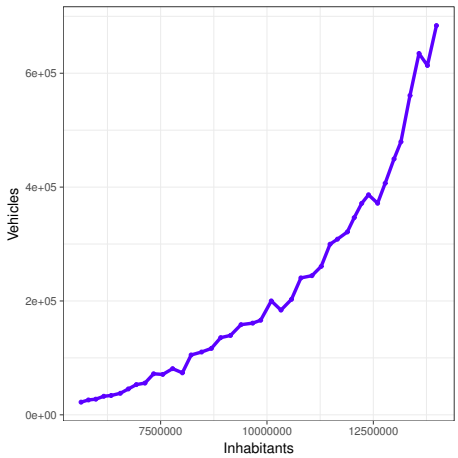


FIGURE 3.2: Ecuador (1968-2008) Inhabitants vs. Vehicles

Traffic congestion is a big issue in DMQ, due to a high concentration of people moving by cars or by bus, the capacity of the infrastructure is not enough for the demand. Besides to congestion, air pollution is another concern to the DMQ's authorities and inhabitants. In a recent report of air quality(Ambiente, 2016), indicates that the main sources of contamination in DMQ are the mobility means as public transportation that uses diesel as a fuel and also the particular vehicles due to high density and congestion. That mobility characteristics described previously and the levels of use of private and public transportation becomes a critical geographical area of study in this research.

The next sections illustrate the area of study and its mobility components.

3.2 Area of Study

The geographical area of study is a large and important part of Quito Metropolitan District (DMQ). The majority of the main facilities: companies, education, hospitals, food markets, public institutions are in that area. It includes the business district, eight major universities, several hospitals, large malls, two large parks, and one major soccer stadium, covering approximately 5 x 8 Km^2 as shown in Fig 3.3.

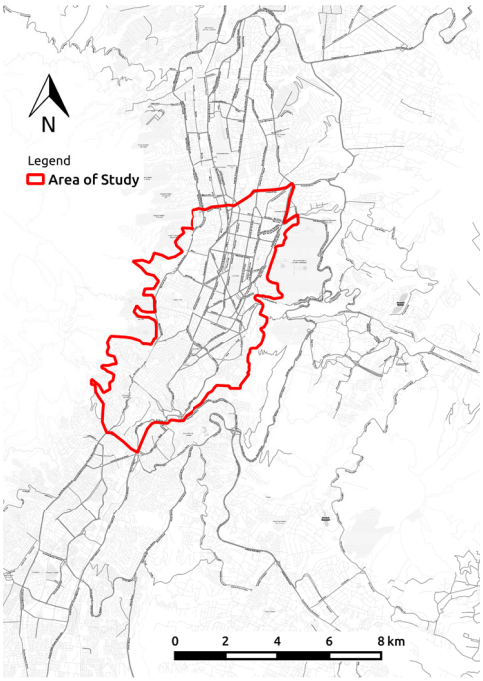


FIGURE 3.3: Area of Study

Moreover, in the city centre, which represents the 17.5% of the total DMQ area has a concentration of the 66% of workplaces that implies a high mobilization to this zone and inside of it. The 90% of the trips go to that zone as a destination from the peripheric zones(Demoraes, 2005).

The next section describes the mobility model components of the area of study: the supply or network infrastructure and the demand.

3.3 How to build DMQScenario

This section describes the method to create the DMQScenario in MATSim. In a first step, as is shown in Fig. 3.4, we build the base scenario configuration files: network, facilities and population. The input data provides the transport network that comes from open sources, activity locations for work and study, and open times for activities given by land use data. The population data provides the neighbourhood density. The fusion process employs a series of libraries and applications to convert to MATSim into XML format.

FIGURE 3.4: Base DMQ Scenario Implementation

3.3.1 Transport Network Infrastructure

The network information is obtained from OpenStreetMap (OSM) as is shown in Fig.3.5. After to select the geographical area and its coordinates, we download the complete geographic database which corresponds to Ecuador from Geofabrik(Frederik, Topf, and Karch, 2007) and extracts the network with the proper coordinates. To extract the network we used an OSM java command line application (*Osmosis*) for processing OSM data. Finally, to convert the network to MATSim XML format, we use a MATSim's network editor application which input is the OSM network file and the geospatial system reference, which for DMQ is EPSG:24877. Fig. 3.6 shows the DMQ scenario used in this research.

The number of links of the network corresponding to the area of study is 8192. The links' attributes are *length* (m), *saturation flow* (capacity) that defines how many cars can pass through the road during a unit of time, e.g. vehicles/hour, and *free speed* that represents the maximum flow velocity. For this work, we take into account all the main and secondary pathways with free speeds equal or above 30 Km/h. Traffic in all main pathways is bidirectional, and some of them include multiple lanes. Traffic in secondary pathways is mostly unidirectional.

Fig.3.7 (a) shows the saturation flow rates of lane groups of the network links. One-lane links are shown in gray and red, two-lane links in blue, green, violet, and orange, and three-lane links in brown. Thus, the saturation flow per lane are between 600 and 1500 passenger cars per hour, which is less than

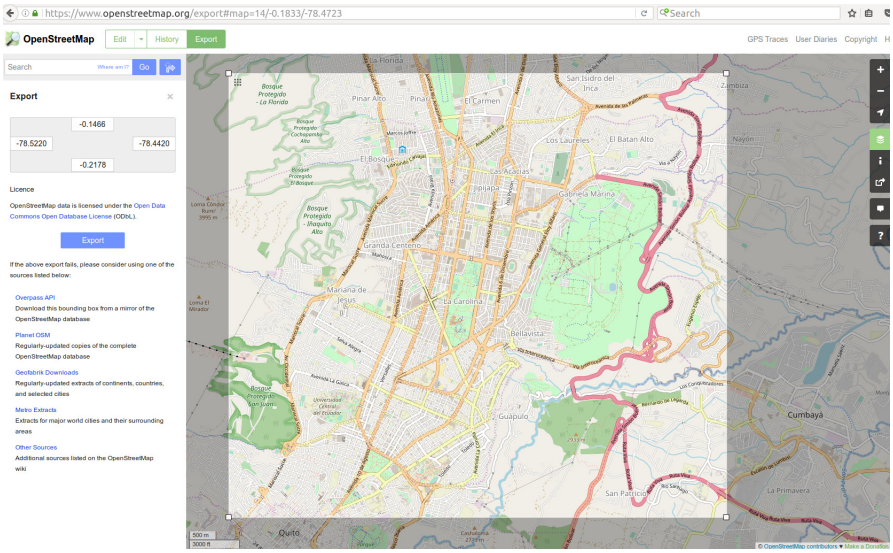


FIGURE 3.5: DMQ from OpenStreetMap

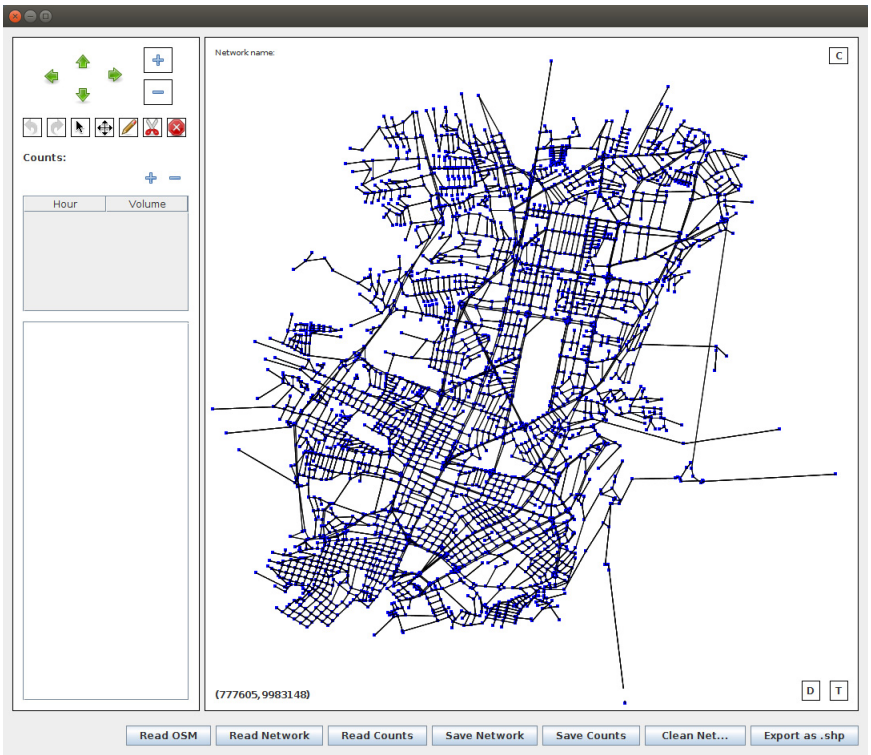
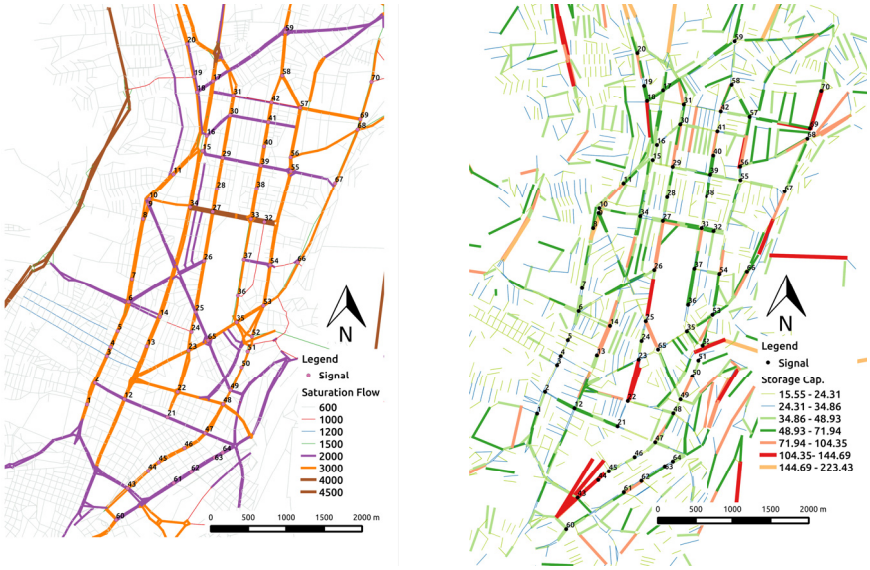


FIGURE 3.6: DMQ Scenario Network

the base saturation flow of 1900 assumed by the The Highway Capacity Manual (Board, 2000). Fig. 3.7(b) shows the storage capacity of network's links which represents the maximum number of vehicles can stand in the queue for each link. MATSim computes the storage capacities internally (it is not a link attribute) by $link_{number-of-lanes} * link_{length}/7.5m$, where 7.5m is the space occupied by a passenger car unit (PCU).



(A) Link's Saturation Flow

(B) Link's Storage Capacity

FIGURE 3.7: Saturation Flow and Network Storage

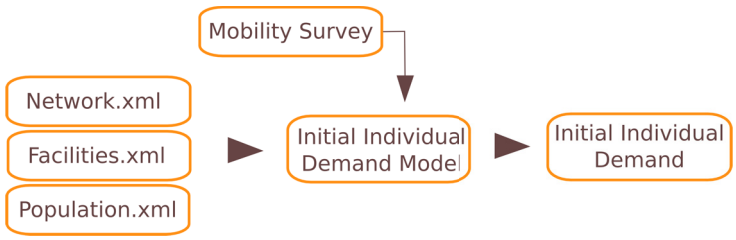


FIGURE 3.8: DMQ Scenario Demand Modeling

3.3.2 Initial Demand - Mobility Model

A second step refers to implement the mobility model. Fig. 3.8 displays the general process to implement the initial demand for DMQScenario. To build the initial demand we use the files described in subsection 3.3.1 and additionally important data that comes from a mobility survey.
Demand is the set of mobility plans per each person. Each mobility plan is defined by some attributes related with activities such as:

- Type: It specifies a kind of activity such as work, study, shop and so on.
- Location: The place which activity is performed.
- Duration: How long does the activity take?.
- Opening or Closing: Time for opening and closing according to each facility.

To simulate the movement of the agents, we need to provide MATSim with their initial mobility plans. In this work, a mobility plan of an agent

consists of two trips. In the first trip, the agent leaves home towards the location of its activity (Castiglione, 2015) and in the second trip, the agent returns home from its activity. Thus, each trip is defined by the coordinates of origin and destination of the trip and the activity specified by the type, starting time and duration of the activity. The type of vehicle used is also associated with the journey.

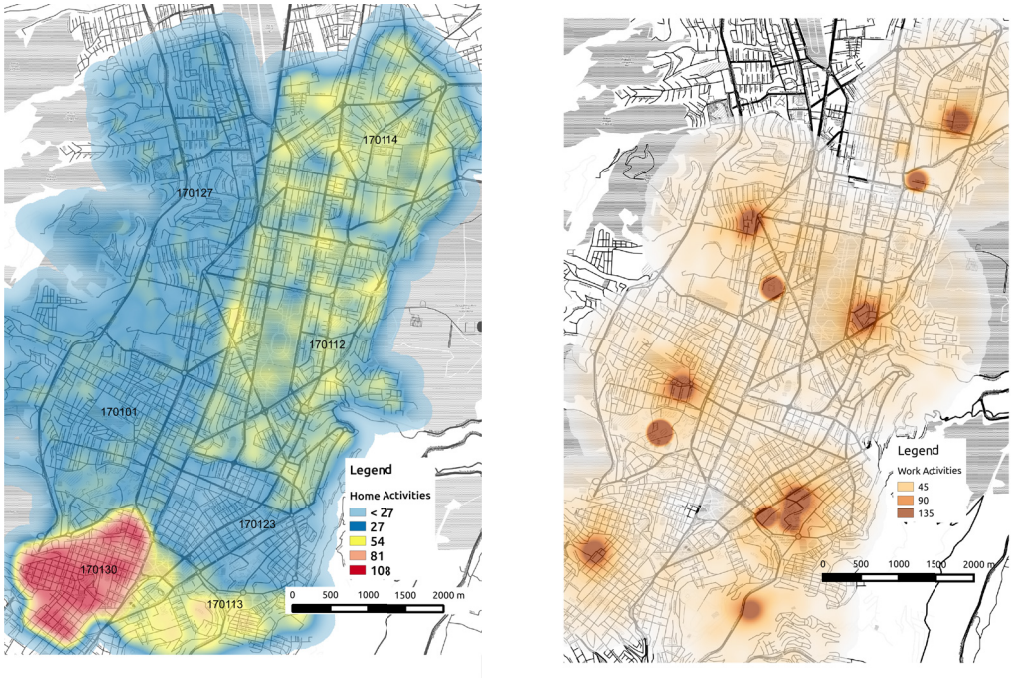
We digitalized the administrative city borders defining a geographical area per district using QGIS Geographic Information System (QGIS Development Team, 2009). We sample home locations for the agents from these districts in proportion to the population according Tab. 3.2.

TABLE 3.2: Scenario Inhabitants per District (2010)

code INEC	District	Inhabitants
170101	Belisario Quevedo	45.370
170112	Iñaquito	44.149
170113	Itchimbia	31.616
170114	Jipijapa	34.677
170123	Mariscal Sucre	12.976
170127	Rumipamba	31.300
170130	San Juan	54.027
	Total	254.115

Location coordinates of activities are determined based on actual information about the city. The origin (home coordinates) of the trips for the agents are determined based on census of the population. Fig.3.9 (a) shows in colors the distribution of the home coordinates of the agents and Tab. 3.2 the population of the districts considered in our simulation. The number of agents from each district is proportional to the district population. Similarly, the destination of the trips (main activity coordinates) are determined based on the location of facilities (companies, universities, malls, and so on) and the fraction of the agents per activity is computed from mobility survey data. We consider three types of activity: *work*, *study*, and *others* for activities such as leisure, business, shopping, and so on. The proportions of agents for these activities are 32%, 33% and 36% respectively, according to mobility survey data (DMQ, 2012). In the case of activity *work*, data about the distribution of workplaces by district in Quito is provided in (Demoraes, 2005). We follow that distribution to assign the coordinates of workplaces probabilistically. For the type of activity *study*, we select probabilistically among the universities located in the area of study and assign its coordinates. The number of agents assigned to each university is proportional to the university's population. In this activity, we neither consider elementary nor high school students. For the type of activity *other*, we use an origin-destination probability matrix to assign the coordinate, where the probability to go to a destination is proportional to the population of the destination district and inverse to the distance. Fig. 3.9 (b) shows the coordinates of the main activities of the mobility plan.

Eq.[3.1–3.6] represents the formulation to obtain the proportion of agents per town in DMQ scenario and Tab.3.3 describes the notation.



(A) Home Distribution

(B) Main Activities Distribution

FIGURE 3.9: Origin and Destination Locations

$$Pop = \sum_{t=1}^{Nt} Pop_t \tag{3.1}$$

$$MovPop_t = \alpha.Pop_t \tag{3.2}$$

$$MovPop_{t,j} = MovPop_t.\omega_j \tag{3.3}$$

$$V_{t,j} = MovPop_{t,j}.\beta_t \tag{3.4}$$

$$PV = \sum_{t=1}^{n}\sum_{j=1}^{m} V_{t,j} \tag{3.5}$$

$$SP_{t,j} = \frac{V_{t,j}}{PV}.SP \tag{3.6}$$

Eq.3.1 computes the total inhabitants of the area of study. In Eq.3.2,variable α corresponds to 80% percent of the total population which moves every day to accomplish several activities.

In Eq.3.3 variable ω represents the fraction value for the main mobility reasons j such as *work, study* and *others* (see table 3.4). As is mentioned before, the value of ω_{study} is adjusted to take into account, only students who are related to high educational level. For that reason, we consider the total university population in DMQ given by the universities located in the zone.

TABLE 3.3: Notation in mobility scenario

Notation	Description
Pop	Total population of inhabitants to model.
Nt	Total number of towns to model.
Pop_t	Total population in t-th town.
α	Population percentage which moves for disparate activities.
ω_j	Mobility factor due to work, study and other activities.
$MovPop_{t,j}$	Population from t-th town which moves due to activity j.
β_t	Vehicles from t-th town per person.
$V_{t,j}$	Vehicles from t-th town per activity j.
PV	Total population of vehicles corresponding to population P.
SP	Syntetic population of agents into simulation.
$SP_{t,j}$	Syntetic population per t-th town per activity j.

ω	Value
ω_{work}	31.1
ω_{study}	32.5
ω_{other}	36.4

TABLE 3.4: DMQ Mobility reasons

In Eq.3.4, the β values come from the ratio of the number of vehicles per family per each town. Eq.3.5 computes the total population by town and activity. Finally, Eq.3.6 computes the final proportion of population per town per activity. The variable SP represents the total number of agents to run in the simulation.

With the final proportion of agents for each town and activity (Eq.3.6), we create a list of agents trips. Important attributes of a trip are defined by the tuple (o, d, a, t_s, t, v), where o is the origin coordinates, d is the destination coordinates, a is the type of activity, t_s is starting time of the activity, t is the duration of the activity, and v is the type of vehicle. Starting times and durations of activities are assigned randomly sampling from ranges defined for each type of activity. Algorithm 5 and 6 create the list of trips for the agents in the simulation.

3.4 Evolutionary Algorithm and MATSim Integration

3.4.1 AnyOEA Framework

AnyOEA Framework comes from Any Objective Evolutionary Algorithms Framework, which is an active project from Shinshu University (Nagano - Japan) and MODO research group (*MODO Frontiers in Massive Optimization and Computational Intelligence*). AnyOEA Framework is a complete set of objects for experimenting and developing single, many, multiple and any objective evolutionary algorithms developed in C++.

Algorithm 5: Initial Mobility Plan

Data: SP,PV,$V_{t,j}$
Result: $\bar{mp} = [origin, activity, dest, starttime, duration, vehtype]$
/* Probabilities for activities destination: Φ_w, Φ_s, Φ_o ,for starting times: $\Phi_{stw}, \Phi_{sts}, \Phi_{sto}$,for duration: $\Phi_{dw}, \Phi_{ds}, \Phi_{do}$,for vehicle type: Φ_v */
$\{\Phi_w, \Phi_s, \Phi_o, \Phi_{stw}, \Phi_{sts}, \Phi_{sto}, \Phi_{dw}, \Phi_{ds}, \Phi_{do}, \Phi_v\} \leftarrow setProbabilities()$;
for t *to* Nt **do**
  $SP_t = \sum_{j=1}^{3} SP_{t,j}$ where $j \in [1 = work, 2 = study, 3 = others]$
  for $i = 1$ *to* SP_t **do**
    if $i <= SP_{t,work}$ **then**
      $act \leftarrow work$;
      $\Phi_{destination} \leftarrow \Phi_w$;
      $\Phi_{starttime} \leftarrow \Phi_{stw}$;
      $\Phi_{duration} \leftarrow \Phi_{dw}$;
    else
      if $i <= SP_{t,work} + SP_{t,study}$ **then**
        $act \leftarrow study$;
        $\Phi_{destination} \leftarrow \Phi_s$;
        $\Phi_{starttime} \leftarrow \Phi_{sts}$;
        $\Phi_{duration} \leftarrow \Phi_{ds}$;
      else
        $act \leftarrow others$;
        $\Phi_{destination} \leftarrow \Phi_o$;
        $\Phi_{starttime} \leftarrow \Phi_{sto}$;
        $\Phi_{duration} \leftarrow \Phi_{do}$;
      end
    end
    $origin = generateCoord(t)$;
    $activity = act$;
    $dest\_town = getDescriptor(\Phi_{destination})$;
    $dest = generateCoord(dest\_town)$;
    $starttime = getDescriptor(\Phi_{starttime})$;
    $duration = getDescriptor(\Phi_{duration})$;
    $vehtype = getDescriptor(\Phi_v)$;
    $\bar{mp} \leftarrow [origin, activity, dest, starttime, duration, vehtype]$;
    $write(\bar{mp})$;
  end
end

Algorithm 6: getDescriptor(Φ)

Data: $\Phi = \{\Phi_i, i = 1...n\}$ where $\Phi_i = \{\Phi_i.descriptor, \Phi_i.weight\}$
Result: Descriptor

```
1  total=0;
2  for i = 0 to |Φ| do
3  |   total=total + Φ_i.weight;
4  end
5  r=random();
6  prob=0;
7  for i = 1 to |Φ| do
8  |   prob = prob + Φ_i.weight / total;
9  |   if r < prob then
10 |   |   return Φ_i.descriptor;
11 end
```

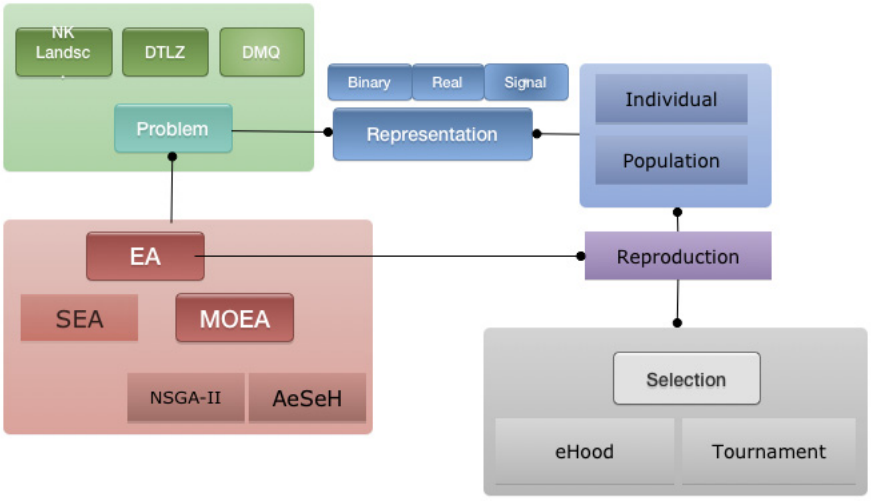


FIGURE 3.10: Evolutionary Algorithm Framework

Fig. 3.10 shows the structure of *AnyOEA Framework*. The main component are:

- EA: Implements the evolutionary algorithm's main flow structure which allows a generic implementation that can be re-used for each specific EA implementation e.g. Single Objective Algorithm, AϵSϵH(Aguirre, Oyama, and K., 2013), etc.
- Problem: Implements a problem behaviour given by the decision variables, the objectives and the way to evaluate each solution.
- Representation: Implements different representations according to decision variables types such as real, integers, binary or problem-oriented solution's codification.
- Reproduction: Implements the variation operators.

- Selection: Implements the selection mode e.g. binary tournament, ϵHood, etc.

3.4.2 MATSim Process

Matsim internally optimizes the demand by a systematic relaxation process until to reach an equilibrium state by the modification of some specific demand variables as route, departure time, activity duration or mobility means. For more details, please refer to subsection 2.4.1.

3.4.3 AnyOEA-MATSim Integration

The integration between simulator(s) and the optimization framework is an operative component of this research. A first step involves communication between MATSim in Java and optimization framework in C++. A first approach using Java Native Interface (JNI) allows bidirectional communication between C++ and Java, which grants to C++ routine to call a method on Java platform. JNI provides a fast communication between Matsim (Java) and AnyOEAFramework (C++). The simulator is used to evaluate each solution from the evolutionary algorithm(AnyOEA). For that purpose, the AnyOEA send the genotype to the simulator and Matsim converts to phenotype and evaluates each solution to get the fitness. Matsim submits the fitness to AnyOEA which uses each fitness to evolves and complete the iteration per each generation.

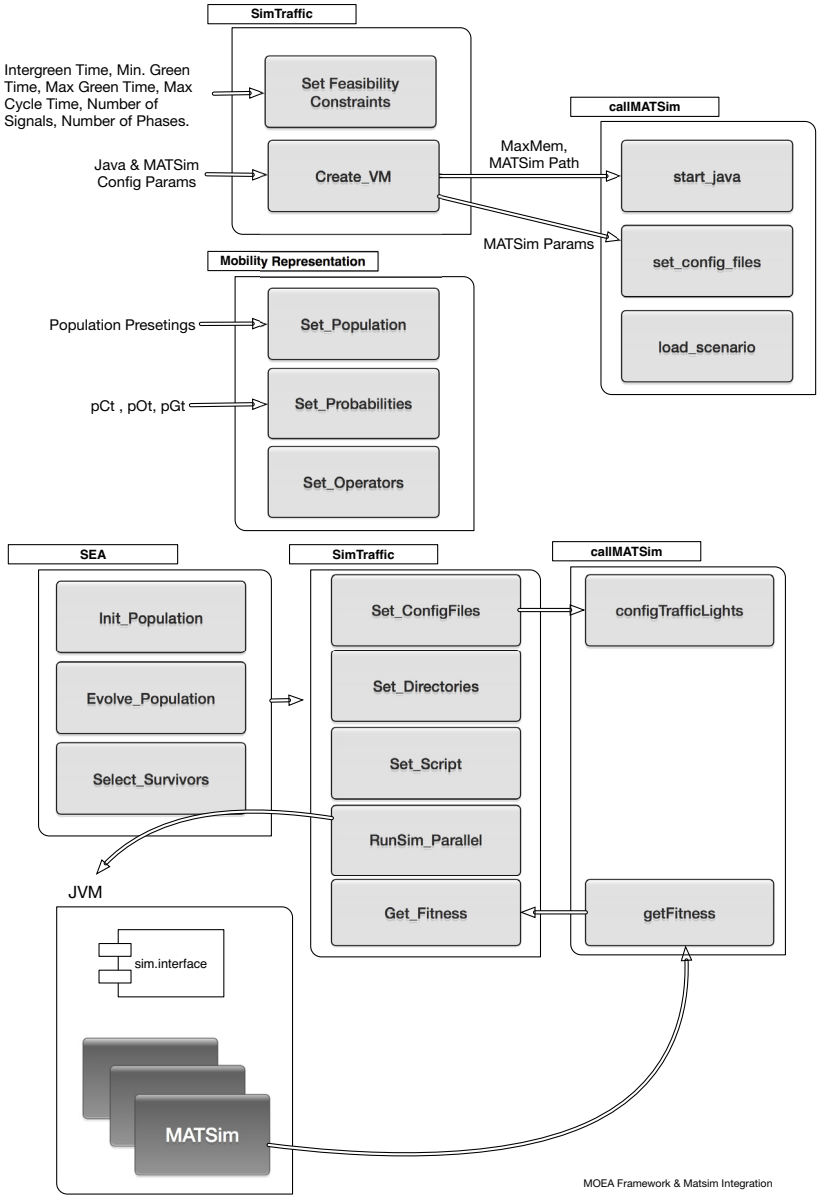


FIGURE 3.11: AnyOEA-MATSim Integration

Fig. 3.11 shows the diagram of the main implemented classes for the traffic signal problem optimization. The implemented classes belong to three types following the AnyOEA Framework: EA, Problem, and Representation. The *SEA* implements a single objective evolutionary algorithm extended from *EA* class. The *SimTraffic* class is extended from *Problem*, and it contains the problem definition. *Mobility Representation* is extended from *Representation*, and it implements the problem's solution codification. Finally, The *callMatsim* class implements the interface between AnyOEAFramework with Matsim.

Chapter 4

Traffic Signals Optimization - Single Objective Formulation

In this chapter, the optimization of traffic signals is studied. A first scenario to test the framework since a single-objective approach is studied. A description of the scenario, algorithm and operators is given. A complete and comprehensive study of the results is also included.

4.1 Introduction

The main objectives of signal timing at an intersection are to reduce the average delay of all vehicles and the probability of crashes(Garber Nicholas, 2009). Another important issue is the signal coordination as a method to provide the ability to synchronize multiple intersections to enhance the operation of one or more directional movements in a system. The intent of coordinating traffic signals is to provide smooth flow of traffic along streets and highways to reduce travel times, stops and delay(STM, 2008).

The re-design of road network infrastructure implies substantial costs that are hardly affordable. Hence, a way to reduce traffic congestion is to make better use of the existing road network, which can be achieved in part by a proper set of traffic signals. Furthermore, proper setting of traffic signals also helps reducing emissions and can induce traffic patterns to control speed in sensible areas to increase pedestrian security.

We concentrate on the optimization of a large number of traffic lights deployed on a wide area of the city and study their impact on travel time, emissions and fuel consumption. A proper setting of traffic signals can help to alleviate the traffic congestion with better use of the current infrastructure. Moreover, it is a key component to study other important problems related to mobility in order to improve the sustainability of the transport system.

We use evolutionary algorithms (EA) to explore a significant number of alternative signal settings under various scenarios of mobility. Activity based micro-simulation allows us to model the mobility demand of each person in the scenario and facilitates a detailed analysis of the traffic in the city. However, it is computationally expensive. This imposes limits on the number of

fitness evaluations, population size, number of generations, and on the number of times the stochastic evolutionary algorithm can be run. In addition, the search space is vast due to the number of signals considered and the several parameters that define a traffic signal. Hence, an important aim of this work is to develop an evolutionary algorithm to search effectively in large decision spaces under a small budget of iterations (generations), performing a reliable short-term evolution to find high-quality solutions. To achieve this, we design a set of specialized mutation operators to search for clusters of coordinated signals with similar cycle length by propagating cycle length between neighboring signals and setting offsets based on the distance between them. Due to the topography of the city and its mobility patterns, it is important to define neighborhoods as two-dimensional and optimize in both directions of traffic flow (north-south-north and east-west-east). Also, we introduce varying mutation rates with high selection pressure to accelerate convergence of the algorithm. The appropriate combination of operators, selection, and varying mutation rates allows to search effectively even with small populations. We use machine learning (ML) for parameter analysis of the evolutionary algorithm to compare our settings with ones suggested by ML methods.

To gain knowledge, we use data mining methods to perform hierarchical clustering in decision space of the best solutions found by the evolutionary algorithm. We include the analysis of signal clusters and their geolocation, estimation of fuel consumption, spatial analysis of emissions, and analysis of signal coordination. This gives an overall picture of the systemic effects of the optimization process.

We verify the effectiveness of the developed algorithm for short-term evolution using a small population. We also show that the design optimization approach is a useful tool to advance the understanding of the transport and mobility problems of Quito city, essential for decision making.

4.2 Related Work

Several techniques have been proposed for signal timing optimization. These range from statistical based methods in the 60's (Webster, 1958; Miller, 1963) to computational intelligence based methods in the last years (Zhao, Dai, and Zhang, 2012) oriented to implement intelligent transportation systems. In the following, we review the literature where meta-heuristics and biologically inspired algorithms have been combined with traffic simulators for optimization. The review is brief and tries to give a broad perspective, but it does not mean that is comprehensive.

The level of traffic simulation has been either macroscopic (Ceylan and Bell, 2004; Rouphail, Park, and Sacks, 2000; Teklu, 2006; Taniguchi and Shimamoto, 2004), mesoscopic (Park, Messer, and Urbanik, 1999), or microscopic (Sánchez, Galán, and Rubio, 2010; Sánchez, Galán, and Rubio, 2008; Garcia-Nieto, Alba, and Olivera, 2012; Garcia-Nieto, Olivera, and Alba, 2013). Some works model toy or virtual scenarios mostly for proofs of concept to

verify the hypothesis about mobility and transport issues, and to test optimization methods (Park, Messer, and Urbanik, 1999; Hong et al., 1999; Ceylan and Bell, 2004; Turky et al., 2009; Chen and Xu, 2006; Tomforde et al., 2008; Peng et al., 2009; Kachroudi and Bhouri, 2009). Other works model real world problems (Sánchez, Galán, and Rubio, 2010; Sánchez, Galán, and Rubio, 2008; Rouphail, Park, and Sacks, 2000; Teklu, 2006) and have mainly focused on small areas of interest with relatively few signalized intersections. For example, seven signalized intersections are optimized in (Sánchez, Galán, and Rubio, 2010) and nine in (Rouphail, Park, and Sacks, 2000). In these studies, usually, the phases of the signals have been modeled in detail, including several phases for some of the signals. However, typically one common cycle has been considered for all signals and some studies do not consider offsets between them. In (Sánchez, Galán, and Rubio, 2010) different cycle lengths are implicitly modeled, but no offsets are considered. Examples of real-world problems with a relatively large number of signalized intersections are (Garcia-Nieto, Alba, and Olivera, 2012; Garcia-Nieto, Olivera, and Alba, 2013), (Teklu, 2006) and (Stevanovic et al., 2011). In (Garcia-Nieto, Olivera, and Alba, 2013) twenty to forty signal controls were considered in an area of 0.42 Km^2, signals were modeled to allow different cycle lengths with two or more phases, but offsets were not explicitly considered. In (Teklu, 2006) seventy-five signalized intersections with two or more phases and offsets between signals were considered in an area of 30 Km^2. However, a common cycle was used for all signals. In (Stevanovic et al., 2011) seventy consecutive signalized intersections located on a 13-mile corridor were optimized. The modeling of the signals is detailed, including 12 NEMA movements and right turns, basic and Transit Signal Priority (TSP) timing parameters, with more than two phases in some intersections. However, it uses a binary coding representation for integers applying one-point crossover and bit flipping mutation. Unfortunatelly, with this representation the magnitude of change depends more on the position of the bit being mutated than on the mutation probability. Thus, tuning of the algorithm and robustness to parameters settings become a serious issue. Even more so when mutation rates are controlled over time, as we do it in our work.

In the above-referenced works, Genetic algorithms (GA) have been the favored optimization technique. Another popular optimization technique has been Particle Swarm Optimization(PSO) (Chen and Xu, 2006; Peng et al., 2009; Kachroudi and Bhouri, 2009; Garcia-Nieto, Olivera, and Alba, 2013). Although most works have focused on fixed settings of the signals, there are also important works on dynamic settings, for example (Hong et al., 1999; Srinivasan, Choy, and Cheu, 2006; Tomforde et al., 2008; Turky et al., 2009; McKenney and White, 2013).

In our work we focus on a real-world network that covers a wide area of 40 Km^2 and include in its main backbone seventy signalized intersections, a relatively large number similar to (Teklu, 2006) and (Stevanovic et al., 2011). We consider a different offset and a different cycle length per signal, instead of a common cycle for all signals. The traffic simulation is agent-based at a microscopic level, and mobility is modeled based on agents' activities, which

allows us to build and test real world scenarios for the city under study. Our study also includes analyses of solutions to understand the systemic effects of the optimization based on several indicators that measure the sustainability of transport systems (Richardson, 2005).

4.3 Problem Formulation

We follow the design optimization approach described in section 2.2 in Fig 2.1. In this work, the objective is to minimize travel time optimizing settings of 70 signalized intersections located in the business centre of Quito aiming to reduce traffic congestion, fuel consumption and emissions. We concentrate on the optimization of a large number of traffic lights deployed on a wide area of the city and study their impact on travel time, emissions and fuel consumption.
We simulate the mobility of a large number of agents and study the impact of traffic light settings on travel time, fuel consumption and emissions. Besides, we analyze optimal solutions to get patterns of signal settings and verify these patterns according to their geolocation. Most signalized intersections in the area of study are two-phased and implement a fixed-time control. Only a few main intersections use left turn signals because most intersections where left turns are allowed have been gradually redesign to include diverging and merging ramps. Also, there are exclusive lanes for some bus lines, but there are no signals with special phases to establish bus transit priority. Recently there are efforts to study the incorporation of some adaptive signals. Since several sectors of the city are highly congested at certain hours, it is anticipated that adaptive signals would be only partially effective. Thus, its is important to have fixed-time optimal schemes that can be used as default timing plans to improve traffic conditions in the city. In Quito city, mobility represents a constant challenge due to the transportation infrastructure is not adequate to the demand.

The principal components of a traffic signal are cycle length, phase, offset, green and inter-green time. *Cycle length* is the time in seconds required for one complete color sequence of the signal. A *phase* is the set of movements that can take place simultaneously. An *Offset* is the time lapse in seconds between the beginning of a corresponding green phase at an intersection and the beginning of a corresponding green phase at the next intersection. In this work we extend the representation used in (Teklu, 2006) to include a cycle per signal. Also, we use integer instead of binary representation. The variables per signal to optimize are cycle length, offset, and green times. We choose these three variables because we aim to find appropriate green times for groups of signals, where signals in the same group share a similar cycle length and are coordinated with the offset but different groups use different cycle lengths. Thus, a solution x with the specification of all n signals considered in the system is represented by

$$x = (S_1, \cdots, S_h, \cdots, S_n), \tag{4.1}$$

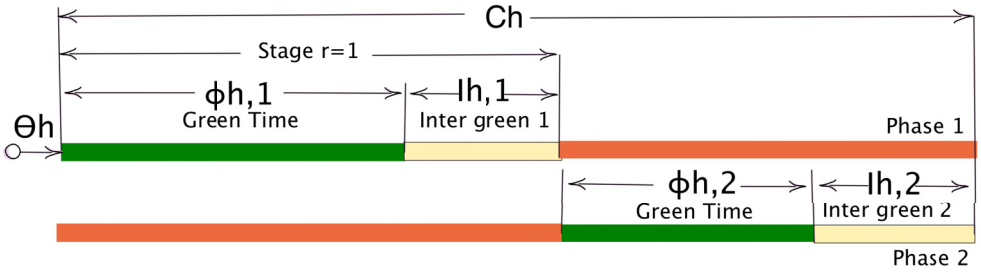


FIGURE 4.1: Traffic light components

where the h-th traffic signal is defined by the following tuple of integer variables

$$S_h = \left(C_h, \theta_h, \phi_{h,1}, \cdots, \phi_{h,r_h}\right). \tag{4.2}$$

Here, C_h is cycle length, θ_h is the offset, and $\phi_{h,1}, \cdots, \phi_{h,r_h}$ are the green times for the r_h phases of the signal. Fig 4.1 illustrates the main components of a signal.

The ranges and constraints of these variables are given in Eq. (4.3) – Eq. (4.9), where $I_{h,j}$ is the inter-green time at signal h for phase j and r_h is the total number of phases of signal h. Equations Eq. (4.3) – Eq. (4.5) represent the range for cycle length C_h, offset θ_h and green time $\phi_{h,j}$, respectively. C_{min} is determined by identifying the signal that needs the longest duration just to accommodate the inter-green times and the minimum green times as shown in Eq. (4.6). C_{max} is set to 135 seconds and inter-green per phase to 3 seconds. These values imply that the minimum cycle time C_{min} is 40 seconds in two phase signals.

The minimum green time should allow drivers to react to the start of the green interval and meet driver expectancy. We follow the Traffic Signal Timing Manual (TSTM)(Federal Highway Administration, 2010) guidelines taking into consideration the driver expectancy(Ellis, 1972) and pedestrian crossing time. TSTM suggests a minimum green time to satisfy driver expectancy between 7 to 15 seconds for major arterials with speed limit of 64 Km/h or less. In our scenario the free speed in most major arterials is around that value. As we mentioned in 4.3 section, the business district has several facilities and services with a high pedestrian volume in some areas. Thus, considering both, the driver expectancy and walk interval duration for pedestrian crossing, we decided to set the minimum green time to 17 seconds for all signalized intersections as shown in Eq. (4.7).

Regarding the maximum green time, TSTM suggests between 40 to 60 seconds for the kind of arterials considered in our scenario. In our work, we search simultaneously cycle length, green times and offsets. As mentioned above the range allowed for cycle length is between 40 and 135 seconds. The maximum cycle length (135) allows us to set maximum green times (as suggest by TSTM) to both phases.

Eq. (4.8) ensures that the sum of the green times in a signal together with inter-green do not exceed the cycle length set for the signal. Eq. (4.9) establishes the maximum green time for the signal phase based on the cycle time, inter-green and minimum green time.

$$C_{min} \leq C_h \leq C_{max} \tag{4.3}$$

$$0 \leq \theta_h \leq C_h - 1 \tag{4.4}$$

$$\phi_{h,j_{min}} \leq \phi_{h,j} \leq \phi_{h,j_{max}} \tag{4.5}$$

$$C_{min} = \max_{h=1,2...,n} \left\{ \left(\sum_{j=1}^{r_h} \phi_{h,j_{min}} + \sum_{j=1}^{r_h} I_{h,j} \right) \right\} \tag{4.6}$$

$$\phi_{h,j_{min}} = 17 \quad sec \quad \forall h, j \tag{4.7}$$

$$C_h = \sum_{j=1}^{r_h} \phi_{h,j} + \sum_{j=1}^{r_h} I_{h,j} \quad \forall h \tag{4.8}$$

$$\phi_{h,j_{max}} = C_h - \sum_{j=1}^{r_h} I_{h,j} - \sum_{k=1,k\neq j}^{r_h} \phi_{h,k_{min}} \tag{4.9}$$

4.4 Simulation and EA Integration

Fig 4.2 shows the integration between the simulator and the evolutionary algorithm (EA). The EA searches optimal settings for the traffic lights. To evaluate a solution, MATSim simulates the mobility of all agents starting from the equilibrium state computed previously and setting its signals controls with the tentative solution provided by the optimizer. MATSim simulates traffic lights microscopically using fixed-time controls (Grether and Neumann, 2011). The output collected from that iteration of the simulator is used to calculate travel time and passed back to the optimizer as the fitness of the solution.

4.5 Evolutionary Algorithm

The optimizer is an evolutionary algorithm described in section 2.5.4 that combines elitism with varying mutations. In the following, we detail the representation, operators, the varying mutation schedule, and fitness function.

4.5.1 Traffic Signals Representation

Fig 4.3 ilustrates the representation of a solution to a system with n signals, each one with r_h phases.

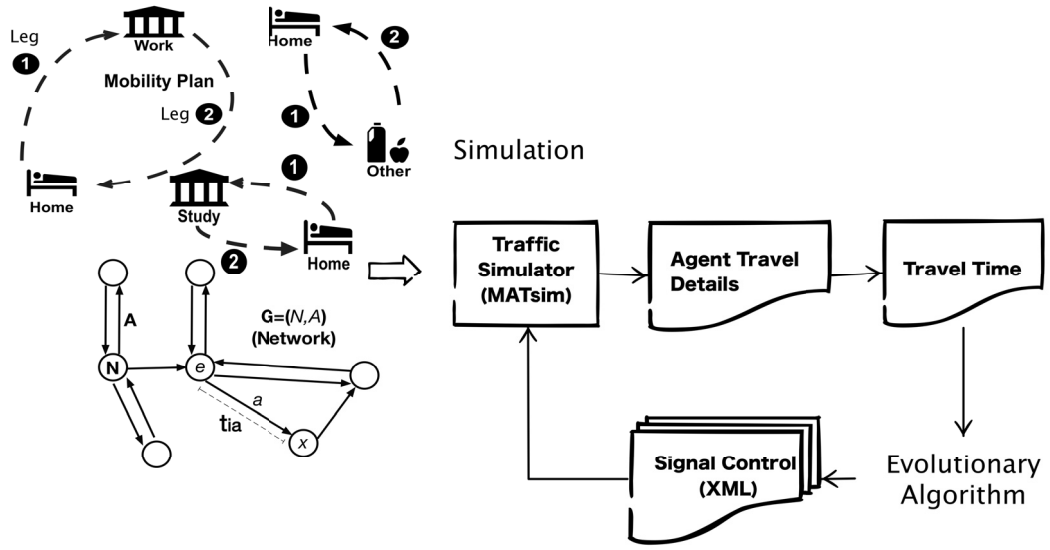


FIGURE 4.2: Simulation and EA integration

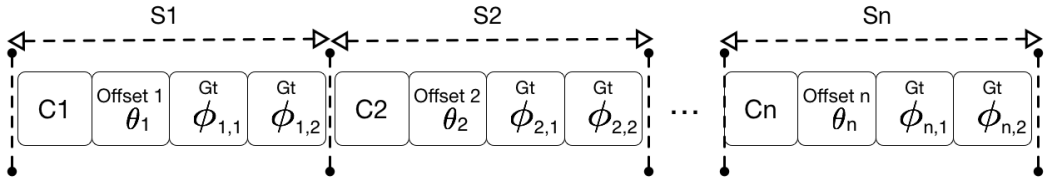


FIGURE 4.3: Solution representation

4.5.2 Operators of Variation

In this work, we use either one or two point crossover to interchanges signals between parents. The crossing points are selected randomly with equal probability in the range [1,*n* - 1], where $\boldsymbol{n}$ is the number of signals.

The algorithm selects between two operators to mutate signals. One is the *Green time mutator (GtM)* and the other one is *Neighborhood cycle propagation with distance-based offset mutator (NCtOPM)*.
The *GtM* operator decreases the green time of one phase and adds it to another phase using step size $\boldsymbol{stepGt}$. The phase $\boldsymbol{i}$ to decrease its green time is randomly chosen among those where the decrement does not violate the constraint for minimum green time $\boldsymbol{\phi_{h,i_{min}}}$. Similarly, the phase $\boldsymbol{j \neq i}$ to increase its green time is also randomly chosen.

The *NCtOPM* operator aims to improve traffic flow along the two main axis of circulation, South-North-South (SNS) and West-East-West (WEW), favoring traffic signal coordination by simultaneously modifying the parameters of a signal $\boldsymbol{S_h}$ and its neighbors $\boldsymbol{N_h}$. The operator stochastically selects the axis of circulation, either NSN or WEW. Then, it chooses one direction of the axis, for example, NS for axis NSN, and propagates the cycle length ($\boldsymbol{C_h}$) of the reference signal ($\boldsymbol{S_h}$) to its neighbors in that direction. In addition, it sets the offsets of the neighboring signals based on the time required to cover the distance $\boldsymbol{d}$ from $\boldsymbol{S_h}$ to the neighbor traveling at free speed. The

operator does the same with the other direction of the axis, i.e. SN. The propagation of cycle and offset to the neighborhood is illustrated in Fig 4.4. Since the cycle length of the signals may change, green times are also adjusted so that the ratios of green time per phase to cycle are the same before and after propagation.

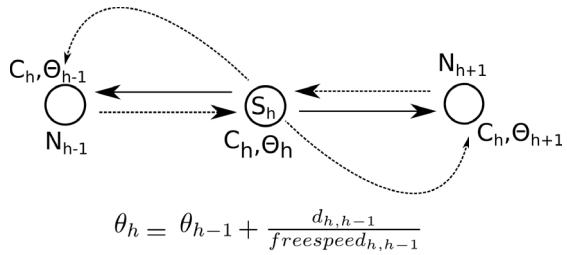


FIGURE 4.4: Cycle and offset propagation

4.5.3 Varying mutation schedule

Varying mutation operators combined with high selection pressure have been proved effective to accelerate the convergence of evolutionary algorithms (Back and Schutz, 1996; Aguirre and Tanaka, 2004). The proposed algorithm introduces elitism inducing a high selection pressure. Also, it includes a time-dependent schedule that deterministically varies mutation rate $P_m^{(t)}$ in the range $[P_m^{(0)}, P_m^{(T)}]$. $P_m^{(t)}$ varies in a hyperbolic shape (Back and Schutz, 1996) and is expressed by

$$P_m{}^{(t)} = \left(\frac{1}{P_m{}^{(0)}} + \frac{\frac{1}{P_m{}^{(T)}} - \frac{1}{P_m{}^{(0)}}}{T-1} t \right)^{-1} \tag{4.10}$$

where T is the maximum number of generations, $t \in \{0, 1, ..., T-1\}$ is the current generation, $P_m{}^{(0)}$ and $P_m{}^{(T)}$ are the desired mutation probabilities per signal at time 0 and T, respectively.

4.5.4 Fitness Function

The fitness of a solution is the average travel time of all agents computed from MATSim's simulation, setting traffic lights with the values encoded in the solution. MATSim outputs, for all agents, the time it takes to travel each link of an agent's route. The fitness function is expressed by

$$\textbf{min} \quad \frac{\sum_{i=1}^{V} \sum_{l=1}^{L} t_{il}}{V} \quad , \tag{4.11}$$

where t_{il} is the travel time on link l for vehicle i, V is the number of vehicles being simulated, and L is the number of links in the network.

TABLE 4.1: Probability of mutation operators P_{CtO} and P_{Gt} and mutation probability P_m and $P_m^{(t)}$ per signal.

Exp	Op.Prob P_{CtO}, P_{Gt}	Mutation Probability per Signal	Observations
E3	0.7,0.3	$P_m = \mathbf{4/70}$	Constant
E4	0.3,0.7	$P_m = \mathbf{4/70}$	Constant
E4DVM	0.3,0.7	$P_m^{(t)} = [P_m^{(0)} = \mathbf{20/70}, P_m^{(T)} = \mathbf{4/70}]$	Varying

4.6 Experimental Setup

4.6.1 Evolutionary Algorithm

We run the algorithm ten times per experiment, use each time a different random seed but always start from the same initial population.We configure three experiments named E3, E4, and E4DVM to test different strategies with mutation operators. The experiments are described with more detail in section 4.7.1.

The number of generations is set to 50, population size is 20, and the number of elite individuals is $\boldsymbol{nElite}$=10. Unless stated otherwise, crossover rate is set to $P_c = \mathbf{1.0}$ and the range for varying mutation per signal $P_m^{(t)}$ is $[\mathbf{20/}n, \mathbf{4/}n]$, where $n = \mathbf{70}$ is the number of signals. To compare with varying mutation probability per signal, we also apply the mutation operators with constant probability $P_m = \mathbf{4/}n$ per signal. Probabilities P_{CtO} and P_{Gt} of the mutation operators $\boldsymbol{GtM}$ and $\boldsymbol{NCtOPM}$ are detailed in Table 4.1 together with the mutation probabilities per signal for the three experiments.

Step size for the green time mutation operator $\boldsymbol{GtM}$ is $\boldsymbol{stepGt} = \mathbf{3}$. In the case of the neighborhood mutation operator $\boldsymbol{NCtOPM}$, we define in advance two neighborhoods for each signal S_h, one for the axis of circulation SNS and another one for the axis WEW. The neighborhoods are based on the actual geographical location of the signals. Also, we pre-calculate the distance and the average free speed to its neighbors along each axis. In this work, the radius of the neighborhood is set to one, i.e. the neighborhood includes the next and previous signal to the reference signal. The neighborhood operator selects for mutation the axis NSN and WEW with probabilities 0.85 and 0.15, respectively, in agreement with the most common traffic flow and mobility patterns in the city.

4.6.2 Initial Mobility Plans

In this work, we design three different mobility scenarios as illustrated in Fig 4.5. In the first one (*S124h*), the trips start from 06:00h and there are several small peaks during the whole day where 500 vehicles or less are in route. In the second one (*S2M*), the agents move during morning hours (06:00h-09:00h) from home to their activity destinations, with a high peak between 07h30 and 08h30 where more than 4000 vehicles are in route. In the last one (*S2A*), the same agents of *S2M* move back home. In S2A the trips

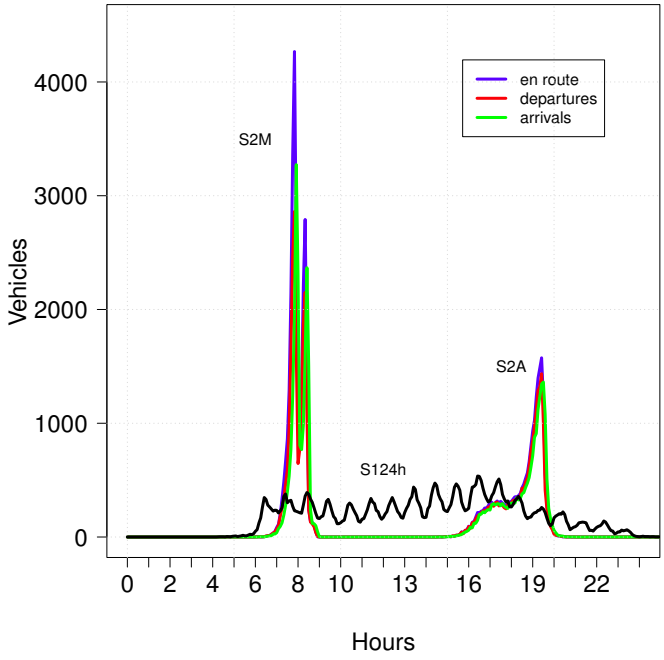


FIGURE 4.5: Mobility Scenarios S124h, S2M and S2A

are spread over a longer period from 15h00 to 20h00 with a peak around 19h00, where approximately 1500 vehicles are in route. These mobility plans of *S2M* and *S2A* are conceived to study the system under saturated scenarios during morning and afternoon peak hours. The three scenarios simulate the mobility of 20.000 agents. Thus, the less saturated scenario is *S124h*, and the most saturated one is *S2M*. This number of agents represents approximately the **30%** of the estimated number of vehicles according to the inhabitants in the zone of study(DMQ, 2012). We assign a type of vehicle to each agent as explained in section 4.7.5. MATSim requires around 8 hours to reach the equilibrium state of 20.000 agents and around 2.5 minutes per individual to compute its fitness. The average number of vehicles is around 3,660 vehicles/hour in the morning rush hours from 7:00 to 9:00 am and 3,655 vehicles/hour in the afternoon hours from 15:00 to 20:00 according to traffic count data taken in 8 intersections located in the area of study. As indicated above, our scenario ***S2M*** simulates a saturated situation that approaches the counters observed in the morning rush hour. The other two scenarios model less saturated situations to test signal settings under different conditions.

4.7 Simulation Results and Discussion

4.7.1 Effects of Operators

In this section, we study the effects of the operators of variation. First, we fix one point crossover and look at mutation operators applied with constant or varying mutation probability.

TABLE 4.2: Mean travel time, standar deviation and interquartile range of best solutions in experiment E3,E4 and E4DVM.

Exp.	Travel Time		
	avg	std	iqr
E3	776.44	30.64	34.04
E4	749.34	20.36	26.53
E4DVM	**746.50**	**17.43**	**21.20**

Fig 4.6 shows the mean travel time over the generations for mobility scenario S124h. In addition, variance of the best solutions at the last generation is shown in detail in Fig 4.7 and Table 4.2 includes numerical values of standard deviation and inter-quartile range.

Experiments E3 and E4 apply the $\boldsymbol{NCtOPM}$ and $\boldsymbol{GtM}$ operators with constant mutation probability $\boldsymbol{P_m = 4/n}$ per signal. E3 applies more often $\boldsymbol{NCtOPM}$ to propagate cycle and offset than $\boldsymbol{GtM}$ to mutate green time. Conversely, E4 applies more often $\boldsymbol{GtM}$ than $\boldsymbol{NCtOPM}$, as shown by the operator probabilities P_{CtO} and P_{Gt} in Table 4.1. Note from Fig 4.6 and Fig 4.7 that mutating more often green times (E4, P_{Gt}=0.7) eliminates outliers and the algorithm converges to travel times lower than propagating more often cycle and offset (E3, P_{CtO}=0.7). Also, from Fig 4.6 note that E3 converges slower than E4. This suggests that configurations with a relatively larger rate for the $\boldsymbol{GtM}$ operator to explore green times combined with lower rates for the $\boldsymbol{NCtOPM}$ operator to propagate cycle length and offsets lead to faster and better convergence.

To study whether larger mutation rates per signal could be useful, in experiment E4DVM we keep the configuration for mutation operators of E4 but instead of using constant mutation rate $\boldsymbol{P_m = 4/n}$ we vary mutation rate in the range $\boldsymbol{P_m^{(t)} = [20/n, 4/n]}$. From Fig 4.6 it is remarkable the fast convergence of the algorithm that applies varying mutations. Note also from Table 4.2 that travel time, standard deviation and interquartile range reduce further compared to E4. We also performed Mann-Whitney-Wilcoxon nonparametric tests, verifying significant statistical differences between E3 and E4 and E3 and E4DVM. Between E4 and E4DVM there is no significant difference in generation 50. This is because both algorithm configurations converge to the same good quality results given an enough number of generations. However, Mann-Whitney-Wilcoxon tests verify that there is a significant difference between generation 25 and 50 for E4 but not for E4DVM, showing that E4DVM converges faster. Fig 4.8 shows the expected number of mutated signals per solution by $\boldsymbol{GtM}$ and $\boldsymbol{NCtOPM}$ when constant and varying mutation are applied.

To verify the search ability and convergence of the algorithm under saturated situations, we run E4 and E4DVM configurations on denser scenarios with larger traffic volume in shorter periods of time compared to S124h.

Fig 4.9 shows mean travel time over the generations on scenarios S2M

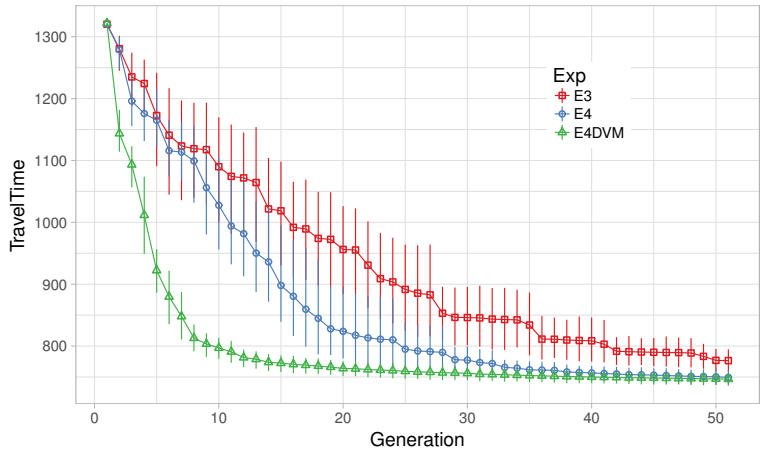


FIGURE 4.6: Mean travel time over generations in experiments E3,E4 and E4DVM

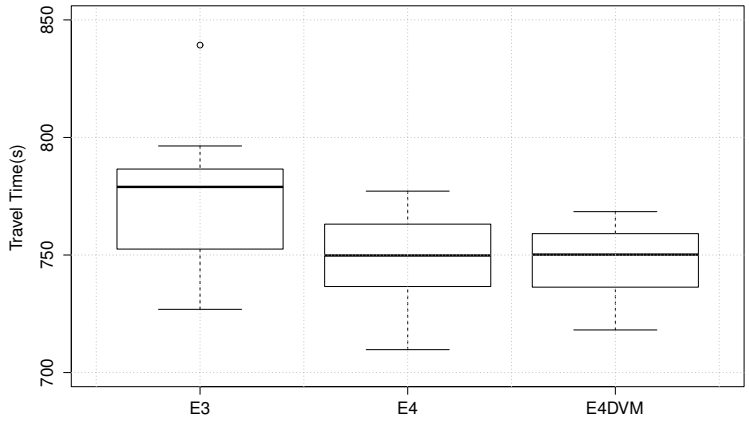


FIGURE 4.7: Mean travel time of best solutions in experiments E3,E4 and E4DVM

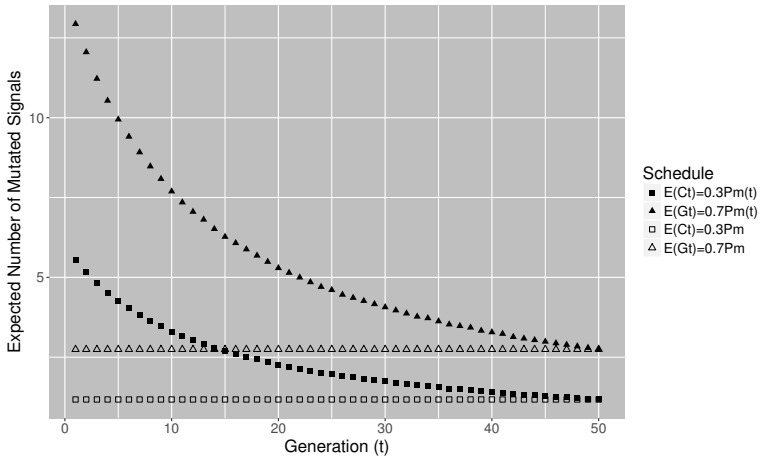


FIGURE 4.8: Expected number of mutated signals with constant and varying mutation schedules (E4,E4DVM)

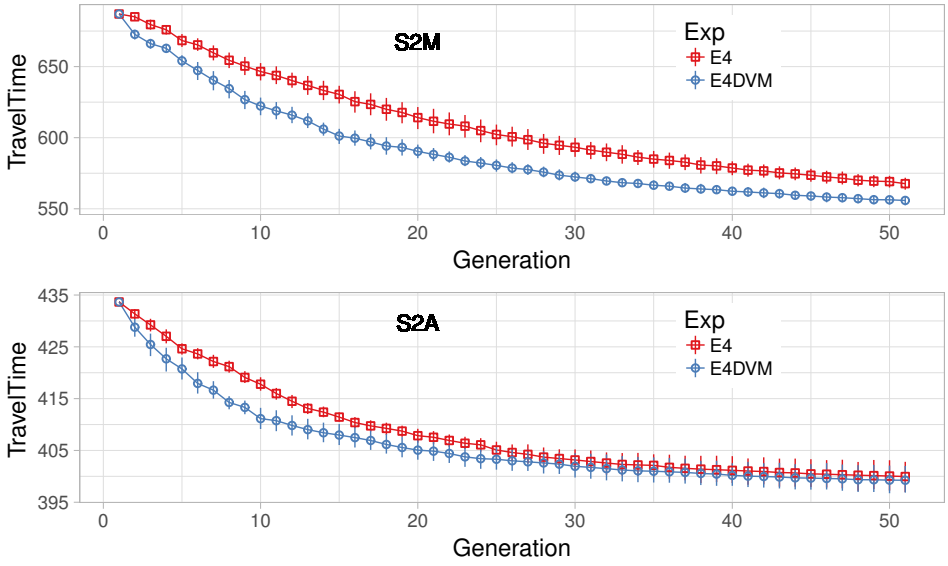


FIGURE 4.9: Mean travel time over generations, scenarios S2M and S2A (E4, E4DVM)

(top) and S2A (down), respectively. From these figures it can be seen that E4DVM convergence faster towards lower travel times than E4. Additionally, it should be noted that the difference in travel time between E4DVM and E4 increases as the scenario becomes denser.

In addition to the design of appropriate mutation operators and their schedule, we also study the effects of crossover. In the above experiments we use one-point crossover with rate $\boldsymbol{P_c = 1}$. We also conduct experiments setting crossover rate to $\boldsymbol{P_c = 0}$ in order to switch off the crossover and replace one-point with two-point crossover keeping the same rate $\boldsymbol{P_c = 1}$.

Fig 4.10 shows mean travel time over generations on the densest scenario S2M by E4DVM with and without crossover. Here, it should be noted that the algorithm using either one- or two-point crossover converges to lower travel time than the algorithm that switch off the crossover. It is known that one-point crossover could be more disruptive than the two-point crossover, particularly when the landscape of the problem is rugged (Spears, 2000; Aguirre and Tanaka, 2003). However, note that in this scenario there is no significant difference between one-point and two-point crossover. In the future, it could be worth studying crossover operators specially tuned for this problem.

The no-crossover configuration is particularly interesting and deserves further discussion. Remember that the same value of cycle length is assigned to all signals of a solution in the initial population. Also, that different values of cycle length are assigned to different solutions. Thus, no-crossover means that signals with different cycle length never get mixed to form a new solution. In addition, the propagation of cycle length between neighboring signals in the same solution by the $\boldsymbol{NCtOPM}$ operator has no effect, precisely because all signals have the same cycle length. In other words, the no-crossover configuration searches only green times by the $\boldsymbol{GtM}$ operator and propagates offsets based on the distance between signals by the $\boldsymbol{NCtOPM}$

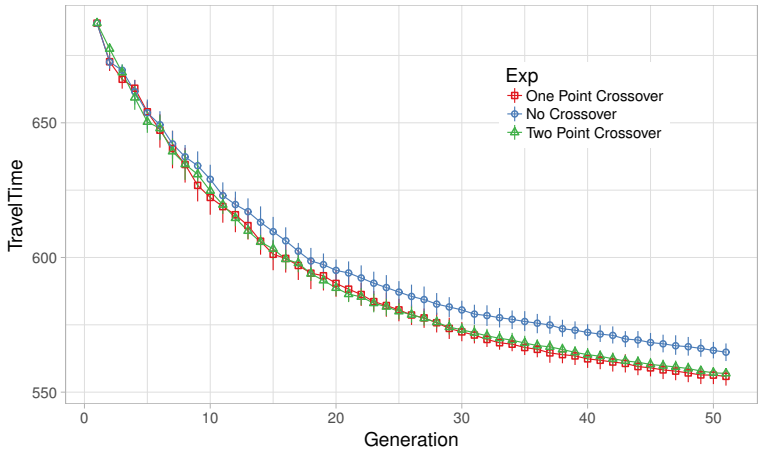


FIGURE 4.10: Effect of crossover operator on mean travel time

operator. This is equivalent to searching an optimal cycle length for all signals instead of one per signal. As can be seen from the figure, better solutions in terms of travel time can be found when we search an optimal cycle length per signal. Other works mostly focus on one common cycle for all signals (Park, Messer, and Urbanik, 1999; Ceylan and Bell, 2004; Teklu, 2006).

The algorithm discussed here is the result of investigating several strategies and configurations of the operators. A first basic approach applies step mutation operators for the cycle and offset, in addition to the step mutation operator for green time discussed here, with an operator probability for each one of them (Armas, Aguirre, and Tanaka, 2014; Armas et al., 2015). A second approach propagates the cycle length and keeps step mutation operators for green time and offset. The introduction of the neighborhood reduces variance and improves convergence speed of the algorithm over the basic approach, however, the search in offsets is not effective with a small number of generations. A third approach, included in this paper, propagates both cycle length and offset in one operator ***NCtOPM*** and keeps step mutation for green time. The inclusion of offset propagation allows an effective search and improves results further. Besides, we also conduct experiments applying step cycle mutation previous to propagate the cycles and offsets in ***NCtOPM***. However, travel time over the generations and final results are almost the same as those obtained when cycle length is propagated without the mutation. These results suggest that the diversity of cycle lengths present in the initial population is appropriate and additional variation of cycles seem not to be required in the scenarios we study here. A detailed discussion and comparison of these approaches using constant mutation probability per signal can be found in (Armas et al., 2016b). As discussed above, another important component of the algorithm is varying mutation. Initial results comparing constant mutation per signal with the varying mutation have been reported in (Armas et al., 2016a).

TABLE 4.3: SMAC solutions: EA parameters

	P_c	$P_m^{(0)}$	$P_m^{(T)}$	P_{CtO}	tt avg.
Conf1	0.60	0.93	0.63	0.19	564.10
Conf2	0.50	0.28	0.61	0.29	554.52
E4DVM	1.00	0.29	0.06	0.3	555.90

TABLE 4.4: Expected Number of Mutated Signals by SMAC configuration and E4DVM

	NCtOPM		**GtM**	
	t=0	t=T	t=0	t=T
Conf1	12.4	8.4	52.7	35.7
Conf2	5.7	12.4	13.9	30.3
E4DVM	6.1	1.3	14.2	2.9

4.7.2 Algorithm's Parameters Analysis

In the previous section, we have shown that varying mutation from high to low rates per signal combined with a strong elitist selection leads to faster and better convergence. Here we use Sequential Model-based Algorithm Configuration (SMAC) (Hutter, Hoos, and Leyton-Brown, 2011) to derive other combinations of parameters settings related to the operators of variation in E4DVM that can lead to good performance. Namely, we analyze optimal combinations of crossover rate P_c, the initial $P_m^{(0)}$ and final mutation probability $P_m^{(T)}$ per signal of the deterministic varying mutation schedule, and the probability of cycle and offset propagation operator P_{CtO}. $\mathbf{1} - P_{CtO}$ determines the probability of the green time mutation operator. Mostly, deterministic varying mutation schedules are used to reduce mutation rates, i.e. $P_m{}^{(0)} > P_m{}^{(T)}$. In our case, we also use it to investigate the effect of increasing mutation rate by allowing $P_m{}^{(0)} < P_m{}^{(T)}$.

Table 4.3 shows the parameter settings Conf1 and Conf2 found by two runs of SMAC using travel time as the observed performance. The settings used by E4DVM are also included for comparison. The settings found by SMAC are different to those used in E4DVM, however, convergence behavior of the algorithm with these configurations and E4DVM are very similar (Armas et al., 2016a). To better interpret the values of these parameters, Table 4.4 shows the expected number of mutated signals per solution at time t=0 and t=T by the two mutation operators in the configurations suggested by SMAC and E4DVM.

From Tables, 4.3 and 4.4 note that the three strategies Conf1, Conf2 and E4DVM mutate green time more often than cycle time, as can be seen by the higher expected numbers of green time mutations compared to the expected numbers of cycle time mutations. Conf2 is a strategy opposite to Conf1 and E4DVM, in the sense that Conf2 increases the expected number of mutations with time whereas Conf1 and E4DVM decrease them. Note that although Conf1 reduces mutations with time similar to E4DVM, the expected number

of mutations is significantly larger in the former. See that in Conf1 at t=0, and t=T the number of expected cycle mutations are two and seven times more than E4DVM, respectively. Similarly, the expected number of green time mutations at t=0 and t=T are three and twelve times more in Conf1 than in E4DVM.

Though Conf1 and Conf2 mutate more than E4DVM, these strategies apply crossover with rate 0.6 and 0.5, respectively. This means that in the configurations found by SMAC, 40 and 50 percent of the offspring are created applying mutation alone (no crossover), whereas in E4DVM the crossover rate is 1.0 and therefore crossover always precedes mutation. It has been shown that the likelihood of interferences between operators increases when mutation with a high rate per variable follows crossover (Aguirre and Tanaka, 2004). This is because a good recombination could be compromised by deleterious mutations or an inefficient recombination could hurt the propagation of beneficial mutations. Thus, smaller rates of crossover combined with higher varying mutations as suggested by SMAC is in accordance with parameter settings that reduce interference between operators. As mentioned above, these three strategies produce similar convergence behavior and lead to a similar low travel time. A more detailed discussion about parameters settings can be found in (Armas et al., 2016a).

4.7.3 Vehicle Movement Distribution

In section 4.6.2, Fig. 4.5 shows the overall vehicle movement distribution over time classifying by departures, arrivals and cars en route. In this section, we focus on the peak hour from 07:30 AM to 08:30 AM to illustrate the vehicle movement distribution in the whole network, in a corridor where 15 signals are located, and with more detail in 3 consecutive signals. The data corresponds to an optimized solution of experiment E4DVM with cycle length of 60 seconds (c60E4DVM).

Fig. 4.11 shows the number of cars per link in the whole network. Links are colored according to the number of cars as shown in the pallete. The numbers close to the black dots identify signalized intersection. Note that in general a larger flow can be observed in the central region than in the perimeter. Also, note that at this hour the largest flow can be seen in the corridor where signals S1 to S6 are located. Fig. 4.12 focuses on the links that belong to the corridor where signals S1 to S15 are located. Here, rows correspond to 15 minutes interval, columns specify links, and colors indicate the number of cars as shown in the pallete.

Similarly, Fig. 4.13 shows the number of cars (nPCU) that crossed a signalized intersection over the time at which green time starts. Results are shown for consecutive signalized intersections S3, S4 and S5 with south-north flow. These intersections are located in a zone close to saturation as shown in Fig.4.11. The storage capacity and max flow capacity per green time of the link where the signal is located are shown as reference in horizontal dotted and dashed lines, respectively. Note that the number of cars counted when they leave the signalized intersections increases from 7:30 to

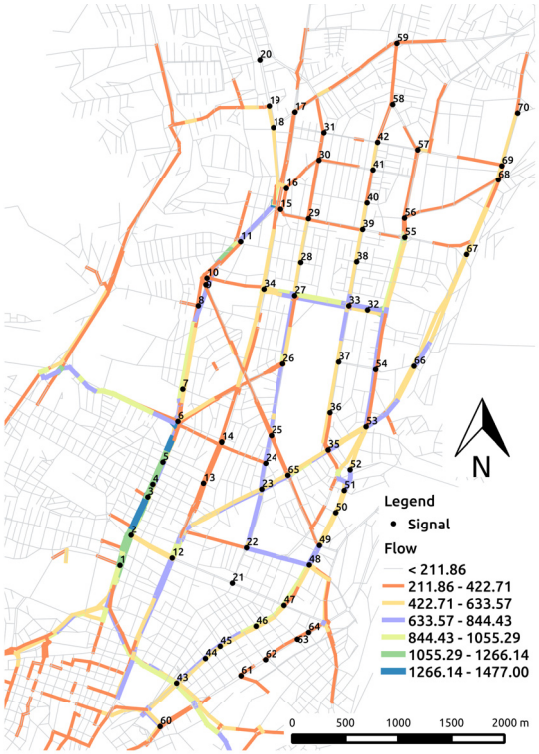


FIGURE 4.11: DMQ scenario flow from 07:30AM to 08:30AM

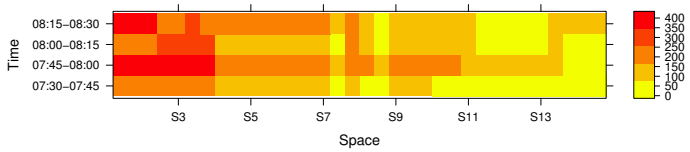


FIGURE 4.12: Signals Corridor S1-S15

7:45 and approaches the max flow capacity per green time around 7:50 to 8:00, which is in accordance with Fig.4.11.

4.7.4 Analysis of Solutions:Decision Space and Cluster Analysis

In this section, we analyze optimal signal settings found by the algorithm.

Fig 4.14 shows the cycle length of the best solutions at the last generation for all runs. Results are shown for experiment E4DVM on the most saturated scenario, S2M. A row corresponds to the settings of all signals in a particular

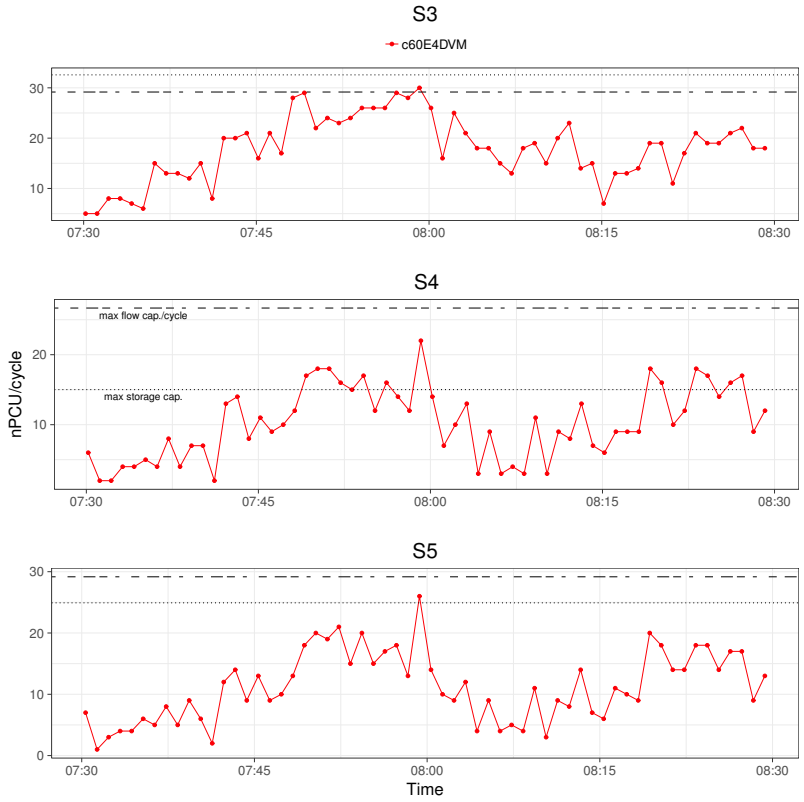


FIGURE 4.13: Signals Corridor S3-S5 flow from 07:30AM to 08:30AM

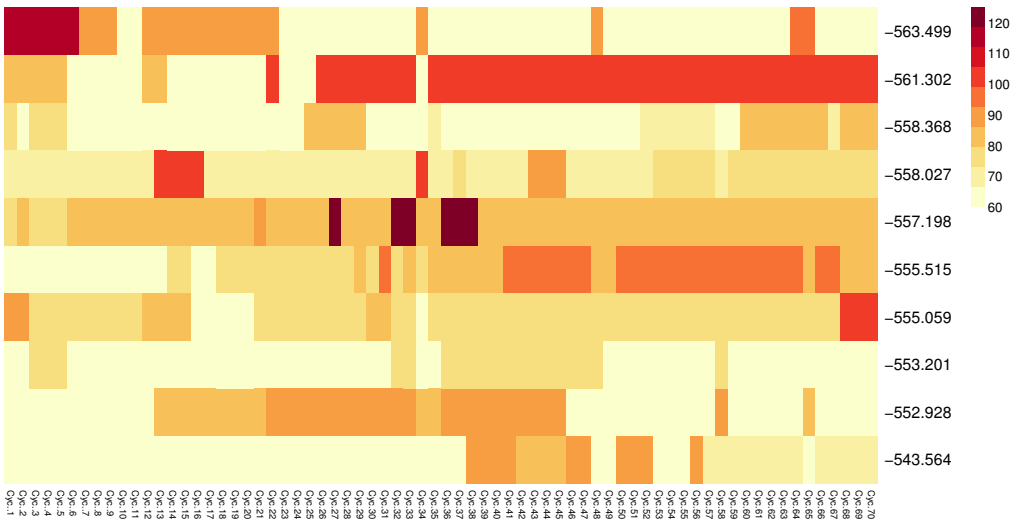


FIGURE 4.14: Cycle length of the best solutions for experiment E4DVM scenario S2M. Annotated heatmap: solutions by row (associated travel time in the row label), variables by column (one per signal), values by color (see color legend).

solution and column identifies a signal. The value of cycle length is shown with color as indicated by the color legend. The labels for each row report the travel time of the corresponding solution.

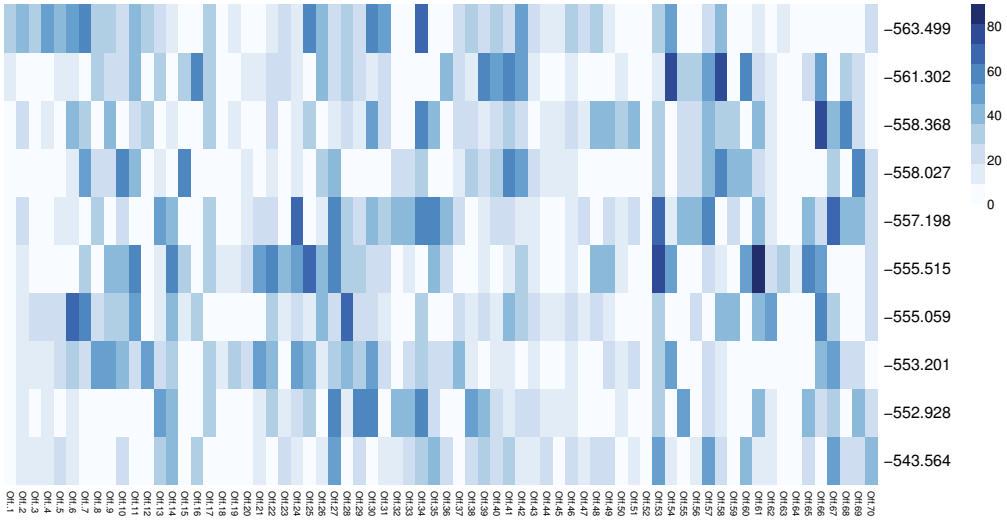


FIGURE 4.15: Offset times of the best solutions for experiment E4DVM scenario S2M

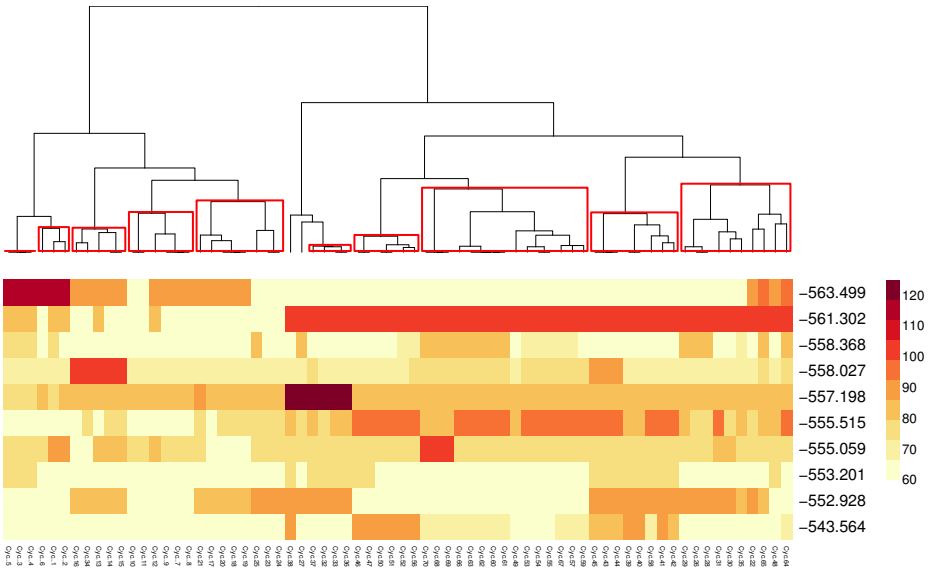


FIGURE 4.16: Hierarchical clustering of cycle length of the best solutions for experiment E4DVM scenario S2M

From this figure, note that groups of signals with the same cycle can be seen in the best solutions found by the algorithm in all runs. This is because the ***NCtOPM*** operator propagates the cycle length value of the reference signal to its neighborhood. However, not all signals have the same cycle. This is an important difference with several approaches that assume one common cycle for all signals (Teklu, 2006) based on simple rules for traffic engineering applicable when the number of signals are few and are closely located (Roess, Prassas, and McShane, 2011). Fig 4.15 shows the values of offsets of the best solutions at the last generation for all runs, similar to Fig 4.14. Note that there is variability of offset values in each solution due to the operator that

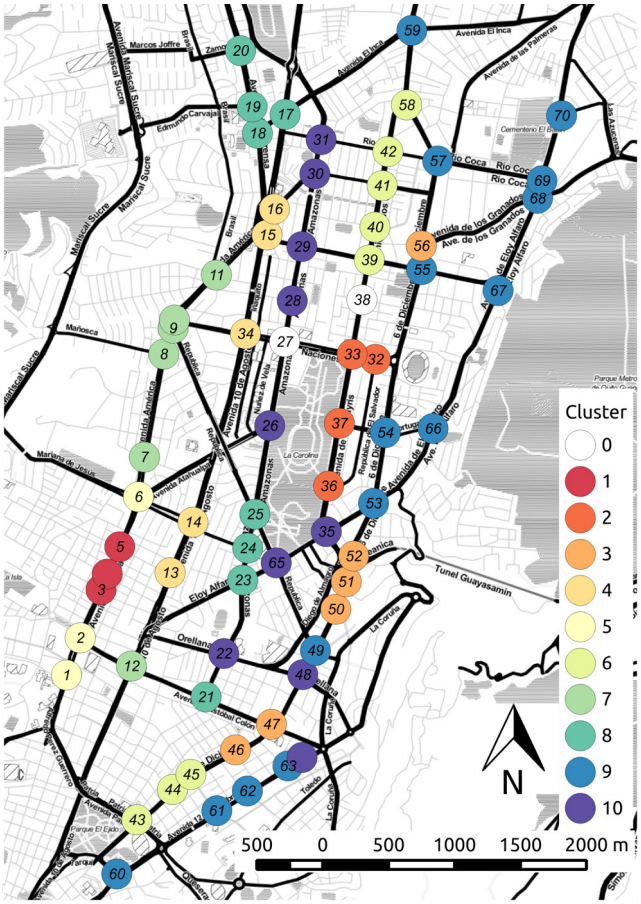


FIGURE 4.17: Geolocation of signal clusters, experiment E4DVM scenario S2M

adjusts the offset based on the distance between neighbor and the reference signal.

An objective of the neighborhood mutation operator ***NCtOPM*** was to induce coordination between signals. Fig 4.14 and Fig 4.15 suggest that there are some patterns in the settings of cycle length and offsets.

To distinguish these patterns and extract some design knowledge from them, we cluster signals based on their similarity. Namely, we apply Ward's agglomerative procedure (Ward, 1963; Murtagh and Legendre, 2014) to create a hierarchy of clusters, using the Euclidean distance between cycle length (or offset) as a measure of similarity. Fig 4.16 shows the hierarchical clustering of cycle lengths represented as an inverse tree diagram, where the dissimilarity between clusters is proportional to the height at which the branches split. That is, the higher the similarity between signals the lower the split. Below the inverse tree diagram, we show the solutions with the columns reordered according to the clustering.

Bootstrapping is used to test the stability of the assignment of variables to clusters. This also gives us a statistically based criterion to determine where to cut the inverted tree and which clusters to keep (Suzuki and Shimodaira, 2015). Clusters that are strongly supported by the data are shown in red in the tree diagram. Signals belonging to the selected clusters are coded by

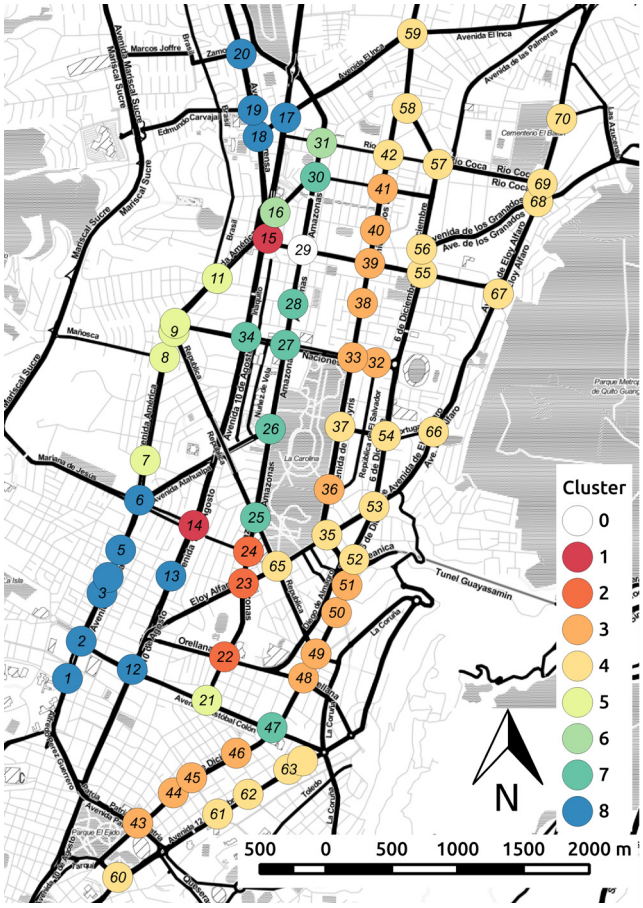


FIGURE 4.18: Geolocation of signal clusters, experiment E4DVM scenario S124h

color and overlaid on the urban map, as shown in Fig 4.17. We can see that most clusters of the signals are aligned with the axis SNS of the principal roads of the transport network. Also, large chunks of the clusters are spatially contiguous. Thus, the cluster geolocalization identifies micro-zones with similar characteristics in the city, which can be further analyzed in order to incorporate additional mobility criteria such as safety, emissions and fuel consumption, multi-modality, etc. Similarly, Fig 4.18 shows the geolocalization of the selected clusters of signals obtained from the best solution of the less saturated experiments E4DVM on S124h. Note that for scenarios under saturated conditions (Fig 4.17) the number of clusters is higher.

4.7.5 Emissions and Fuel Consumption

In this section, we analyze the effects of optimizing traffic light settings on fuel consumption and gas emissions. We use MATSim emission extension (Hülsmann et al., 2011) that computes the gas emissions per link per agent, at the time an agent enters a link of the transport network. It calculates warm and cold-start exhaust emissions for cars by connecting MATSim simulation output to the detailed database Handbook on Emission Factors for Road

Transport (HBEFA) (Keller and Wuthrich, 2014). We have selected seven categories of vehicles based on model year, fuel type and weight. Table 6.1 shows the distribution of vehicle categories chosen for the scenarios, which is in accordance with census transportation data (INEC, 2010b) for Quito city. We assign a category of vehicle to each agent randomly following this distribution.

MATSim computes warm emissions deriving the kinematic characteristics from the simulation and combines this information with vehicle characteristics to extract emission factors from the database of the Handbook on Emission Factors for Road Transport (HBEFA). To derive the kinematic characteristics, the emission model considers 'free flow' and 'stop&go' as traffic states per road segment. To calculate cold-start emissions, MATSim derives parking duration and accumulated traveled distance from the simulation. For parking duration, HBEFA database differentiates emission factors in one hour time steps from 1h to 12h. After 12 hours the vehicle is fully cooled down. There are also different cold emission factors for short trips (less than 1Km) and longer trips (greater than 1Km)(Kickhofer, 2014).

TABLE 4.5: Car distribution (fuel=gasoline and weight $\leqq$ 2Tons)

Year-Category	%	Year-Category	%
2000-2002 ǀ Euro2	24.40	2008-2010 ǀ Euro5	13.70
2002-2004 ǀ Euro3	7.80	2010-2012 ǀ Euro6	23.80
2004-2006 ǀ Euro4	12.20	2012-2014 ǀ Euro6	5.20
2006-2008 ǀ Euro4	13.00		

Table 4.6 shows travel time (TT) and fuel consumption (FC) together with HC, CO, NOx, and $\boldsymbol{CO_2}$ emissions produced by all agents on S124h. First, we show as reference results corresponding to the equilibrium state without traffic signals and the solution with smallest travel time at generation 0 when traffic signals are included. Note that at generation 0 all signals are set with the same cycle length, as explained before. Next, we show results for solutions with traffic signals that minimize travel time at generation 50 in experiments E4 and E4DVM. Comparing with generation 0, results at generation 50 illustrate that in addition to minimizing travel time, fuel consumption and the various kinds of emissions can also be reduced significantly if traffic lights are optimized. Further, the experiments E4 and E4DVM represents an incrementally better optimization process for TT, as discussed above. Note that FC and emissions also reduce incrementally in these experiments. Thus, a better optimization process for TT, given by the neighborhood operators and deterministic varying mutation, is also highly correlated to a better optimization for FC and emissions.

Similarly, Table 4.7 shows results for experiments E4 and E4DVM with one point (XP1) and two point crossover (XP2) on the most saturated scenario S2M. Note from this table that significant reductions in FC and emissions can also be obtained in this kind of saturated scenarios.

TABLE 4.6: Emissions of best solutions on scenario S124h

	Eq.St.	g=0 C_h=130	g=50 E4	E4DVM
TT	591.28	1302.06	694.17	700.19
FC	14275.90	15201.26	14961.03	**14864.17**
CO2	26960.42	29150.52	28434.11	**28348.46**
PM	0.28	0.29	0.29	0.29
NOx	32.41	33.38	33.06	**33.01**
NO2	1.08	1.14	1.12	1.12
SO2	0.13	0.14	0.14	0.14
CO	772.98	774.13	773.96	**773.91**

FC in Liters - Emission in Kg. - TT in s.

TABLE 4.7: Emissions of best solutions on scenario S2M.

	Eq.St.	g=0 C_h=85	E4	g=50 E4DVM(XP1)	E4DVM(XP2)
TT	386.63	650.62	530.96	517.98	524.11
FC	6356.91	6883.16	6771.98	**6746.11**	**6746.52**
CO2	14903.61	16553.44	16204.88	**16123.79**	**16125.08**
PM	0.15	0.16	0.16	0.16	0.16
NOx	15.75	16.55	16.40	16.36	16.36
NO2	0.55	0.59	0.58	0.58	0.58
SO2	0.07	0.08	0.08	0.08	0.08
CO	482.93	484.14	484.09	484.01	483.99

FC in Liters - Emission in Kg. - TT in s.

4.7.6 Spatial Analysis for CO_2 Emissions

In the following, we analyze carbon dioxide emissions ($\boldsymbol{CO_2}$) on a spatially disaggregated level for scenario S2M. MATSim output emissions are aggregated per link for all agents. For geolocation on the urban area of the scenario, $\boldsymbol{CO_2}$ emissions are spatially smoothed using the inverse distance weighted interpolation method (Shepard, 1968). For that, we used as sampling points the centroid of each link and its corresponding accumulated measure of emissions.

Our spacial analysis of the emissions is not an indication of how the emissions are being dispersed in the atmosphere, instead, we illustrate the origin and intensity of the emissions to compare an optimized traffic signal plan against a not optimized one.

Fig 4.19 shows the difference in $\boldsymbol{CO_2}$ emissions between the best solution in the initial population, where all signals have the same cycle $\boldsymbol{C_l = 85}$, and the best solution at the last generation of experiment E4DVM. We use the single-value cycle length ($\boldsymbol{C_l = 85}$) as reference for comparison to show the

improvement that can be achieved from same-cycle solutions at the initial population and to highlight the relevance of having different cycle lengths for different clusters of signals at the end of evolution.

We use a red-blue palette to show emission differences, where levels of red represent increments and levels of blue represent reductions of emissions. White is used to show no difference in emissions between the two cases. Note the dark blue regions, which show that emissions mostly reduce across the area of study when signals settings are optimized. However, note that in some regions there is a relative increase in emissions. This figure gives an overall view of the effects of optimization in CO_2 emissions. It provides useful information for city planners and can be used to feedback the evolutionary algorithm to favor low emissions in certain regions of interest.

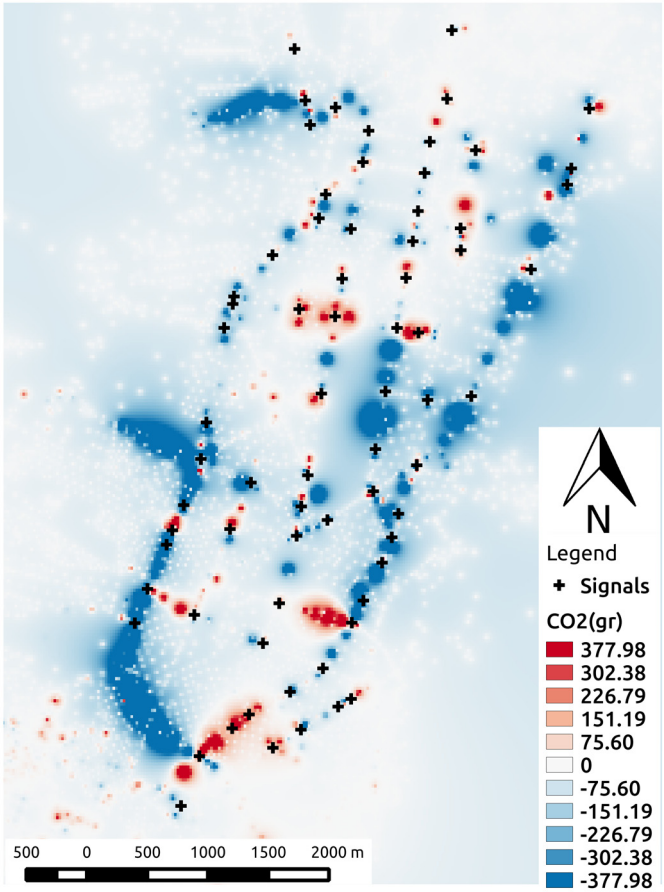


FIGURE 4.19: Change in CO_2 emissions

4.7.7 Coordination Analysis

An important aspect of signal coordination is improving the traffic flows and reduce congestion. A continuous traffic flow during several intersections in one main direction is usually referred as a *green wave*. In this section, we analyze the optimized signal settings to determine if their coordination induce green waves or not.

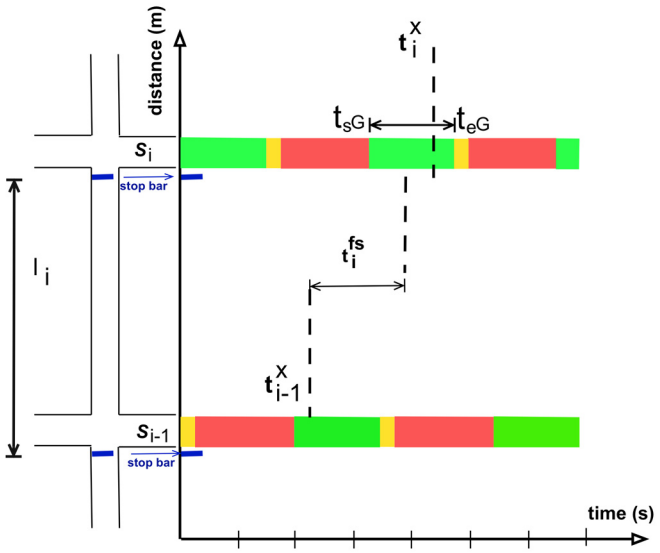


FIGURE 4.20: Variables to detect green phase crossing

Let us denote l_{i-1} and l_i two links with consecutive signals S_{i-1} and S_i located at the end of each one of them, as illustrated in Fig 4.20. The time at which an agent exits link l_i and crosses signal S_i is t_i^x. The starting and ending time of the green time phase when the agent crosses signal S_i are t_{sG} and t_{eG}, respectively.

The earlier expected arrival time to signal S_i after crossing signal S_{i-1} is $t_{i-1}^x + t_i^{fs}$, where t_i^{fs} is the time that takes to cover the distance of link l_i traveling at free speed, i.e. the maximum speed at which an agent can circulate at a given link.

In this work, an agent is said to cross two consecutive signalized intersections in an uninterrupted flow if the earlier expected arrival time to signal S_i and the actual exit time from link l_i (cross signal S_i) are within the green phase window of signal S_i, i.e.

$$t_{sG} \leq t_{i-1}^x + t_i^{fs} \leq t_{eG} \wedge t_{sG} \leq t_i^x \leq t_{eG} \tag{4.12}$$

Fig 4.21 represent the phenotype expression of some consecutive signals of the best individual of experiment E4DVM on the saturated scenario S2M. Fig 4.21(a) shows in different rows the offset and three cycles (green, intergreen, red) of consecutive geolocated signalized intersections s39–s42 and s58 that belong to cluster 6 as shown in Fig 4.17. Similarly, Fig 4.21(b) shows consecutive geolocated signalized intersections s66–s70 that belong to cluster 9. A visual inspection of these signals settings suggests that green waves could emerge from them. To verify this, we count the number of agents that cross these signals.

Table 4.8 and 4.9 show the percentage of agents that cross two or more consecutive green signals. In addition, it shows in detail the number of agents that cross 2, 3, 4 and 5 consecutive signals. Results are shown for

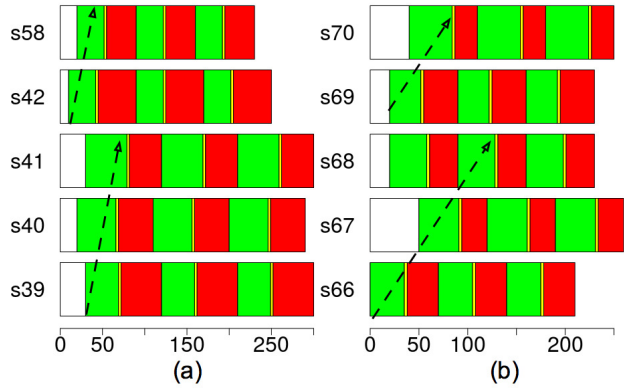


FIGURE 4.21: Best individual (lowest travel time) cluster No.6 (a) and cluster No. 9 (b) for South-North direction, experiment E4DVM, scenario S2M

gw %	C_h85 0.49	E4 0.42	E4DVM 0.47
2	137	123	73
3	20	–	43
4	–	–	14

TABLE 4.8: Green waves cluster 6 total agents=253

gw %	C_h85 0.39	E4 0.45	E4DVM 0.68
2	554	377	556
3	60	230	245
4	–	18	127
5	–	1	61

TABLE 4.9: Green waves cluster 9 total agents=1400

the best solution in the initial population, where cycle length of all signals is $C_h = \mathbf{85}$, and the best solutions of experiments E4 and E4DVM. Note that in cluster 6 there is not much difference in percentage between the initial solution and the optimized ones, although for E4DVM some agents can cross up to 4 consecutive signals. On the other hand in cluster 9 there is a significant increase in percentage in the optimized solutions and some agents can cross up to 4 or 5 consecutive signals. Looking at Fig 4.19 note that in the region where cluster 6 is located there is no reduction in CO_2 emissions. However in the area where cluster 9 is located there is a significant reduction in emissions.

4.8 Conclusions and Future Work

This work presented a design optimization framework for the transportation system of Quito. An evolutionary algorithm to search efficiently using small populations in few generations was proposed. The algorithm was coupled to the multi-agent transport simulator MATSim to study the optimization of a large number of traffic signal controls located on a wide area of the city. Three mobility scenarios of 20,000 agents modeled based on activities were used to verify the effectiveness of the evolutionary algorithm. It was shown that

green times should be mutated with higher probability than cycle and offsets, the propagation of cycle length to neighboring signals with offsets set based on distance leads to better coordination of signals, and the use of varying mutation increases convergence speed. The combination of these strategies efficiently and effectively explore the large search space of traffic signal parameters. The effect of crossover was also verified, showing that better convergence can be achieved when the crossover is activated, although no significant difference between one and two point crossover was observed. The proposed algorithm combines a strong selection pressure given by elitism with crossover followed by varying mutation, where an appropriate balance between exploration and exploitation can be achieve in several ways. An analysis of the parameters of the algorithm using sequential model based algorithm configuration (SMAC) confirmed our finding that mutation of green times should be emphasized over cycle length and offset propagation. It also showed that other configurations with varying mutations relatively higher than the suggested in our configuration but using smaller rates of crossover lead to similarly good results.

Also, hierarchical clustering was performed on the best solutions found in several runs of the algorithm. An analysis of signal clusters and their geolocation, estimation of fuel consumption, spatial analysis of emissions, and an analysis of signal coordination provided an overall picture of the systemic effects of the optimization process. It also showed that the proposed approach is helpful to deepen our understanding of the problem and gain knowledge about the system.

There is ongoing work on multi-objective formulations of the traffic signals problem, where criteria for sustainable transport systems are considered as objectives to be optimized.

In the future, we would like to include the multi-modality of the transport network and study ways to improve the sustainability of Quito's transport and mobility system. Additionally, we would like to include signals with more than two phases. Also, it will be important to combine adaptive signals with default pre-fixed timing plans that are optimal under certain conditions or scenarios.

Chapter 5

Traffic Signals Optimization - Bi Objective Formulation

In this chapter, a bi-objective approach of the traffic signal optimization problem is considered. Here an analysis of the trade-off between travel time and fuel consumption is studied. It concludes with a discussion of the results.

5.1 Introduction

In this work, we integrate MATSim (Horni, Nagel, and Axhausen, 2016), the Comprehensive Modal Emission Model (CMEM) simulator (G.Scora and Barth, 2006), and a multi-objective evolutionary algorithm. The mobility scenario is the same used for the single objective optimization problem described previously, where 20.000 vehicles move and 70 signal lights are located on Quito's business district. We aim to study and understand the influence of optimal signal settings on travel time and fuel consumption. Particularly, we want to clarify the extent of the conflict between these objectives, if any, when they are optimized simultaneously and how the settings of the signals relate to the trade-offs between them. We also compare with a single-optimization algorithm where only travel time is optimized and evaluate the impact of the signals settings on gas emissions.

The three main components of the optimization system considered in this study are MATSim, CMEM simulator, and a multi-objective evolutionary algorithm. Figure 5.1 illustrates their interaction. CMEM is a microscopic emissions simulator that computes second-by-second tailpipe emissions and fuel consumption based on different vehicle operating modes (modal), such as idle, steady-state, cruise, and various levels of acceleration/deceleration. Before we run the optimizer, we prepare the initial mobility plans of the agents as well as the model of the transport infrastructure and run MATSim without signal lights until it reaches an equilibrium state. Also, we prepare the profiles of the vehicles associated with the agents, which are required by CMEM.

The multi-objective evolutionary algorithm evolves a population of candidate solutions. Each solution represents the configuration of all light signals (signal control) of the transportation system under study. The algorithm minimizes simultaneously two fitness functions, the average travel time and the fuel consumption of the agents that move in the transport network. At

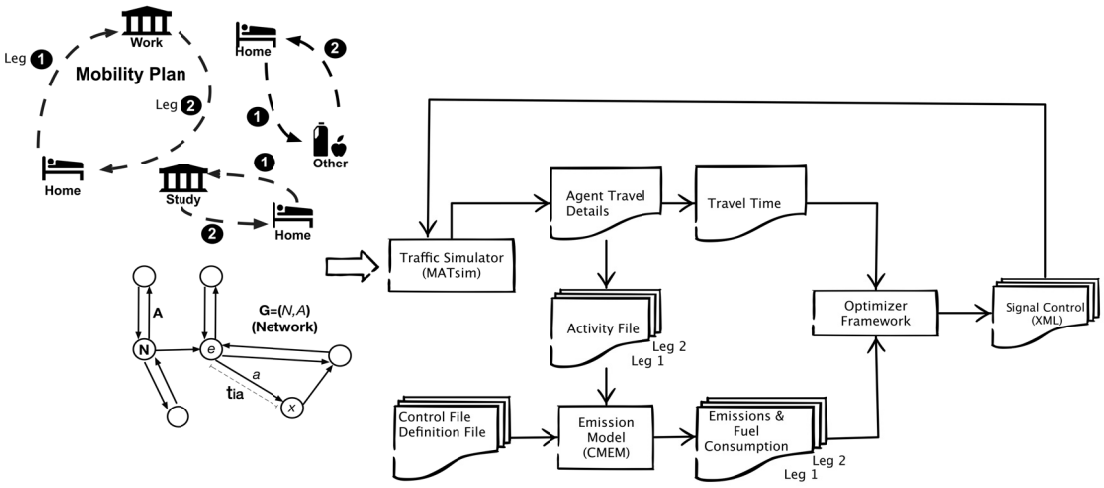


FIGURE 5.1: Optimization System

each generation, to compute the fitness of a solution, the evolutionary algorithm calls MATSim and CMEM, one after the other. MATSim sets the signals of the transport system with the values specified by the tentative solution provided by the evolutionary algorithm. Then, MATSim runs one iteration to simulate the movement of the agents following the mobility plans and routes that led the system to the equilibrium state. The output generated by MATSim is used to compute the average travel time of the agents. CMEM is called with the travel details of each agent extracted from the MATSim output and the profiles of the vehicles prepared in advance. The output generated by CMEM is used to compute the fuel consumption of the agents. Once all solutions are evaluated, the evolutionary algorithm continues to the next generation, stopping after a specified maximum number of generations has been completed.

5.2 Multi-objectivization

Multi-objectivization refers to a method which implies the addition of other objectives to the single objective optimization (SOO) problem to become multi-objective and compare two solution that is common in multi-objective optimization (MOO) techniques using the notion of inferiority under Pareto optimization. In that way, the multi-objectivization method may open up monotonically increasing paths to the global optimum that are not available under the original SOO problem (Knowles, Watson, and Corne, 2001). By the multi-objectivization method, we want to analyze the effect on the solutions when travel time and fuel consumption are optimized simultaneously and compare with a single-optimization algorithm where only travel time is optimized.

5.3 Evolutionary Algorithm

In this work we use the Adaptive ε-Sampling and ε-Hood (AεSεH) (Aguirre, Oyama, and K., 2013) algorithm described in section 2.5.4 to search optimal solutions.

5.3.1 Representation & Operators

For this work, the representation and operators are the same as described in single-objective problem in section 4.5.2.

5.3.2 Fitness Functions

In this work, we minimize two fitness functions, the average travel time and the total fuel consumption of the agents that move in the network. To compute the fitness of a solution, MATSim sets the signals of the system with the values specified by the solution passed by the evolutionary algorithm, simulates the movement of the agents following the routes that led the system to an equilibrium state, and outputs the time taken by each agent to travel each one of the links included in its route. A transport network can be represented by a directed graph $G = (N, A)$, where N represents nodes and A represents links. The travel time for a given vehicle is

$$t_{ia} = t^x_{ia} - t^e_{ia} \quad a = 1, ..., A; \quad i = 1, ..., V, \tag{5.1}$$

where t_{ia} represents the travel time on link a for vehicle i, t^x_{ia} denotes the time vehicle i exited link a (see Figure 5.1), t^e_{ia} denotes the time vehicle i entered link a, V is the number of vehicles being simulated, A is the number of links in network, e is the entry node and x is the exit node (Spiegelman, Sug-Park, and Rilett, 2011). Thus, the average travel time, the first fitness function, is expressed by

$$f_1 = \frac{\sum_{i=1}^{V} \sum_{a=1}^{A} t_{ia}}{V}, \tag{5.2}$$

subject to signal timing design and feasibility constraints shown in Eq. (4.3)-Eq. (4.9) (**Teklu2007**).

The second fitness function corresponds to the fuel consumption of the agents along their legs. It is computed from the output generated by CMEM, which is called along with the travel details of the agents produced by MATSim and the profiles of the vehicles. The second function is stated by

$$f_2 = \sum_{i=1}^{V} \sum_{j=1}^{L} c_i^j \tag{5.3}$$

where V is the number of vehicles, L the number of legs, and c_i^j is the fuel consumption (in grams/km) of the i^{th} vehicle at the j^{th} leg.

5.4 MATSim and CMEM Preliminaries

The scenario includes 70 signal lights located on the main pathways with flows in south-north-south, and east-west-east directions (see Figure 5.5). We run the multi-agent transport simulator MATSim for 200 iterations, making sure it reaches a user equilibrium state without setting any traffic signal. The traffic simulation period is for 24 hours. It takes approximately 10 hours of computation time to run MATSim for this number of iterations. Traffic signals are optimized using the equilibrium state as an initial condition.

CMEM uses a total of 55 static parameters to characterize the vehicle tailpipe emissions for the appropriate vehicle/technology category. CMEM defines 24 Light-Duty Vehicle (LDV) categories based on fuel and emission control technology, accumulated mileage, power to weight ratio, emission certification level, and emitter level category. We have selected 4 categories based on two main features: accumulated mileage and emitter level category based on model year distribution according to transportation census data (INEC, 2010b). Table 5.1 shows the vehicle categories chosen for our scenario. We assign randomly a category to each agent according to the distribution obtained from the census.

TABLE 5.1: CMEM Vehicle Categorization

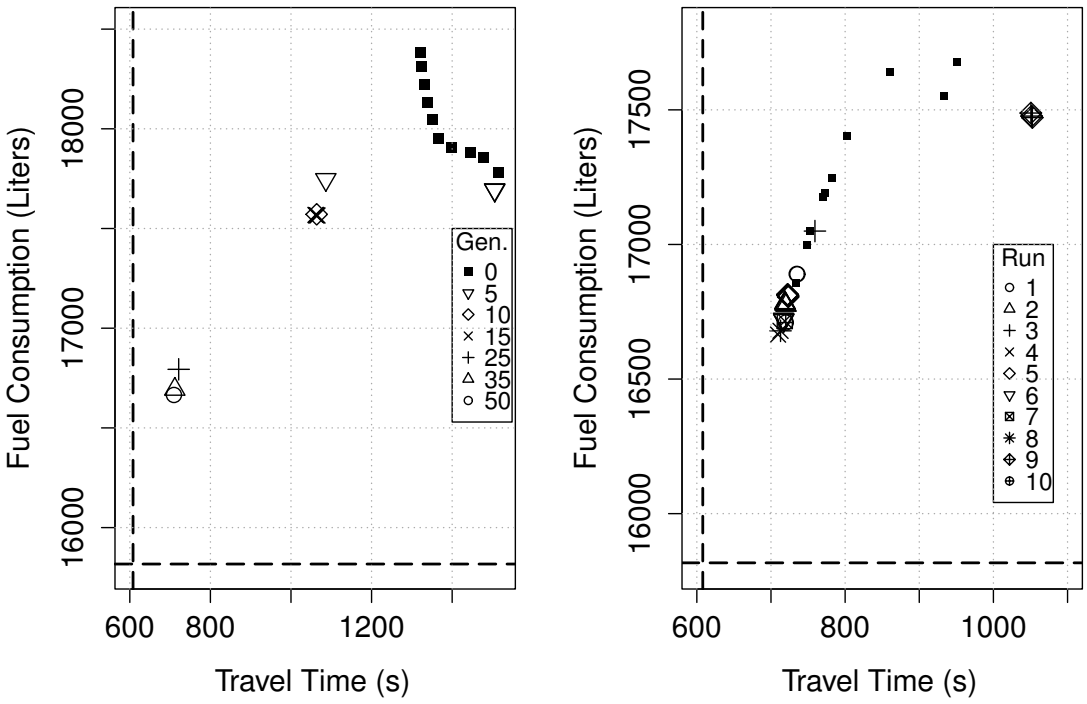
	LDV Categories
9	Tier 1>50K miles high power/weight
24	Tier 1>100K miles
26	Ultra-Low Emission Vehicle
27	Super Ultra-Low Emission Vehicle

5.5 Evolutionary Algorithm Experimental Setup

We use a fixed population size of 20. The initial population is created deterministically as follows. We prepare 20 cycle lengths in the range [40, 135] seconds in steps of 5. All solutions are set with a different cycle length, but all signals of a solution are set to the same cycle length. The offset times of all signals are set to zero and green times per phase are set to the same value according to the cycle length, i.e. green time = (cycle length - inter-green) /2. That is, all signals are synchronized to start at the same time but are not coordinated to allow the uninterrupted flow of vehicles along contiguous signals in the same pathway.

For the operators, we set crossover rate to $P_c = \mathbf{1.0}$ and mutation rate per signal to $P_m = \mathbf{4/}n$, where n is the number of signals. The mutation rates for cycle length, offset and green time operators are ${P_m}^{(Ct)} = \mathbf{0.5}$, ${P_m}^{(Of)} = \mathbf{0.3}$ and ${P_m}^{(Gt)} = \mathbf{0.2}$, respectively. The mutation steps are set to: $stepCt$=5, $stepOff$=10, $stepGt$=3 for cycle, offset and green time, respectively. These mutation steps reduce considerably the search space.

We conduct 10 runs of the algorithm setting the number of generations to 50, use different random seeds but all runs start with the same initial population. To evaluate one individual, it takes in average 4 minutes to run MATSim and compute the first fitness function, and 16 minutes to run CMEM and compute the second fitness function.



(a) Bi-objective optimization (b) Single- and bi-objective optimization

FIGURE 5.2: Objective values of solutions by bi- and single-objective optimization

5.6 Simulation Results and Discussion

Figure 5.2 (a) shows the Pareto fronts found by the algorithm at generations $\{0, 5, 10, 15, 25, 35, 50\}$ for one of the runs. Fuel consumption is converted to liters from kilograms using a gasoline density of 0.755 Kg/liter. The intersection of the dashed lines marks the fitness value of the solution at equilibrium state without signals. Note that a clear trade-off between travel time and fuel consumption can be observed at generation 0 in the initial population. As evolution proceeds, travel time and fuel consumption reduce and approach the values observed at equilibrium state, but the number of non-dominated solutions reduce to a few and in some generations even to one. These results illustrate that the optimization of signals allowing different cycle times and coordinating them by properly setting their offsets lead to significant

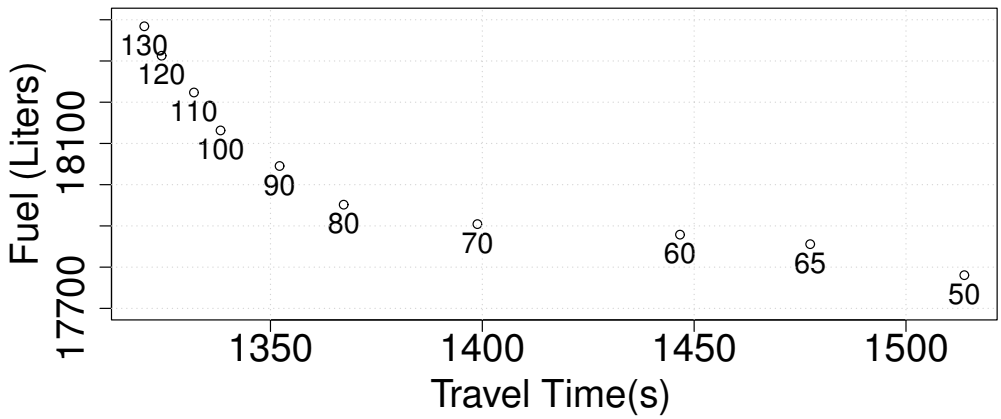


FIGURE 5.3: Non-dominated solutions in initial population labeled with cycle length

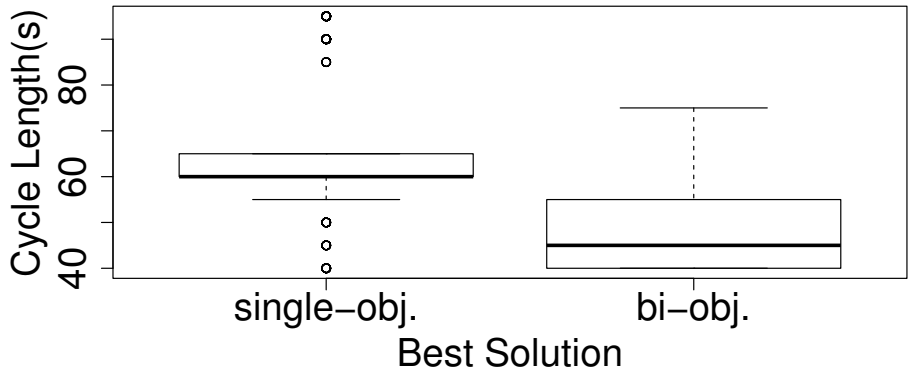


FIGURE 5.4: Cycle length, best travel time solutions, single- and bi-objective optimization

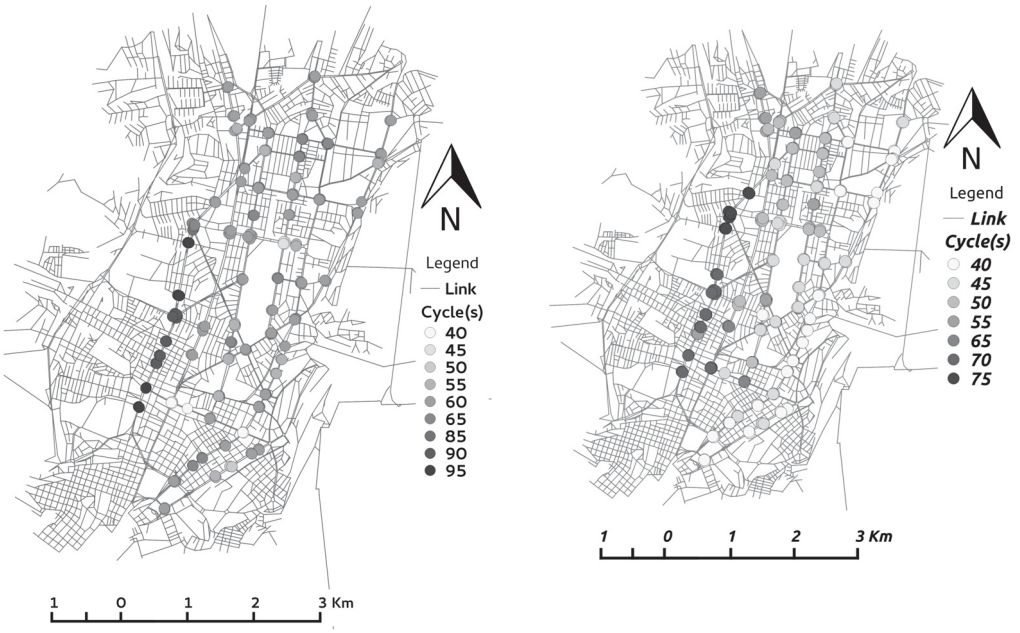
reductions in both fuel consumption and travel time. The small number of non-dominated solutions in the last generation is expected, because both objective functions are correlated. That is, a reduction in travel time implies that the engines are turned-on for a shorter time and therefore use less fuel.

Here an important question is whether optimizing a single fitness function, either travel time or fuel consumption, could be enough to minimize both objectives. We verify this by optimizing only travel time with an elitist single-objective optimization algorithm (Armas, Aguirre, and Tanaka, 2014) set with the same initial population, operators, and parameters used for the bi-objective optimizer. Figure 5.2 (b) shows the Pareto fronts found at the last generation of the 10 runs of the bi-objective optimization algorithm. It also includes in black squares the best solutions found by the single-objective

optimization algorithm. From this figure note that overall results by the bi-objective optimization are better than by the single-objective optimization, thought both point towards the same minimum values. These results suggest that although there could be few non-dominated solutions in the region where both objectives are minimized the inclusion of the second objective helps to perform a more effective optimization. It is also worth noting that variance by the single-objective optimization is larger than by the bi-objective optimization. Nonetheless, the multi-objective optimization could also get trapped in local optima far away from the region of optimality, as observed for run 9 in Figure 5.2 (b) where travel time and fuel consumption are around 300 seconds and 800 liters worse than the solutions with minimum fitness found in run 4. This is a computationally very expensive problem, and not many runs are possible. Thus, it is important to reduce the variance of the solutions found in different runs to increase the reliability of the algorithm. To that end, we should analyze further the operators, population size, and selection of the algorithm in order to find ways to escape local optima.

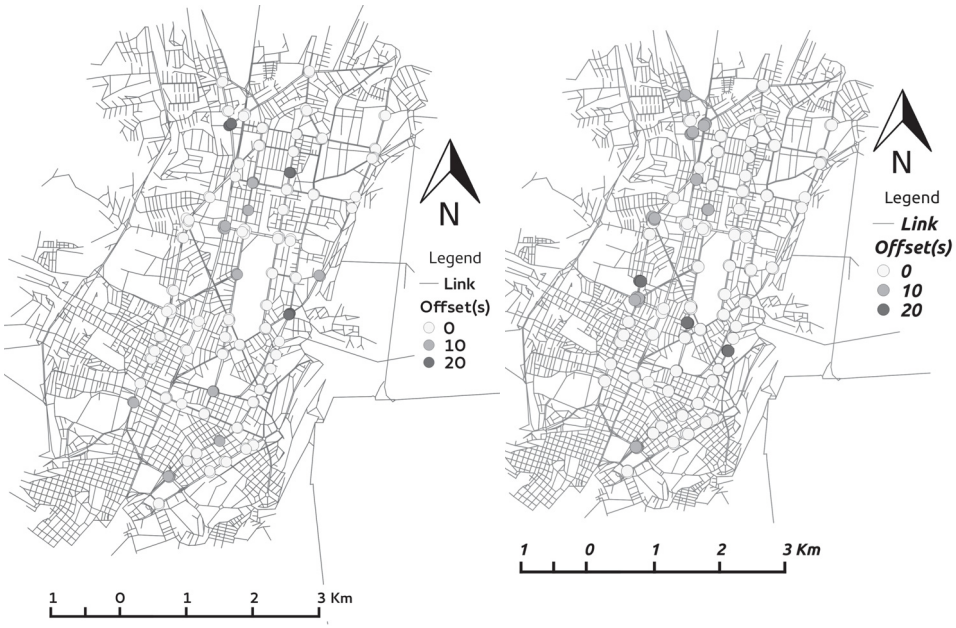
Figure 5.2 (a) and (b) illustrate the trade-offs in objective space. In the following, we analyze the settings in decision space, particularly cycle length and offset of the signals. Figure 5.3 shows the non-dominated solutions in the initial population, fuel consumption over travel time (labeled with cycle length), where all signals of a solution are set to the same cycle length, offset is set to 0, and green times are similar in both traffic flow directions. Note from the figure that when signals are not coordinated, offset set to 0, smaller travel times are achieved by longer cycle lengths and lower fuel consumptions are achieved by shorter cycle lengths. Figure 5.4 shows box-plots of the cycle length of the best solutions in travel time found by the single- and bi-objective optimization. Note that the optimized solutions include shorter cycle lengths than the best solutions in the initial population and that the cycle lengths by the bi-objective optimization are shorter than by the single-objective optimization. For the single-objective algorithm the highest ranked solution are the ones with the larger cycle length. So, those solutions will be preferred for mating and reproduction. This could imply a loss of diversity of solutions with shorter cycle lengths. However, as indicated above, optimal solutions are a combination of signals with shorter but different cycle lengths. In the case of the bi-objective optimizer, solutions with shorter cycle length will also have a high rank thanks to the second objective, i.e. fuel consumption. Thus, the bi-objective optimizer will not suffer from a lack of diversity of solutions with shorter cycle length. This explains why the bi-objective approach performs a more effective optimization than the single objective approach.

Figure 5.5 shows the cycle length of the signal lights of the solutions with shorter travel time by the single and bi-objective optimization approaches, deployed on the map of the area of study. Similarly, Figure 5.6 shows the offsets of the signal lights. From Figure 5.5 it is worth noting that a pattern can be seen in the solutions produced by both approaches. In both solutions, the largest cycle lengths are assigned to signals located in the south-north avenue in the western part of the city. This illustrates the kind of design



(a) Single-objective (b) Bi-objective

FIGURE 5.5: Best Solution Cycle



(a) Single-objective (b) Bi-objective

FIGURE 5.6: Best Solution Offset

knowledge we aim to extract from the optimization process, useful to understand and decide the final settings of the signal lights. From Figure 5.6 it

should be noted that both solutions include some signals with offset 10 or 20, however still many of them remain 0. This is due to the short-term evolution used in this work. The proper setting of offsets undoubtedly helps improve traffic. In the future, we should look for ways to enhance the optimization of offsets.

TABLE 5.2: Scenario Emissions

	Eq. State	g=0 C_h=130	g=0 C_h=50	g=50 Bi-obj.	g=50 Single-obj.
Travel Time (s)	608	1320	1513	709	734
Fuel Consum. (l)	15817	18384	17780	16665	16858
HC (Kg)	121.97	129.76	128.26	125.81	126.32
CO (Kg)	2623.90	2764.50	2762.97	2736.81	2737.88
NO_x (Kg)	277.44	253.71	262.34	264.18	261.89
CO_2 (Kg)	33347.37	39248.40	37810.00	35188.40	35647.50

Table 5.2 shows travel time and fuel consumption together with HC, CO, NO_x, and CO_2 emissions produced by all agents corresponding to the equilibrium state without traffic signals. Also solutions including traffic signals at generation 0 with smallest values in travel time and fuel consumption, and solutions with traffic signals that minimize travel time by the bi- and single-objective optimizer at generation 50. These results illustrate that in addition to minimizing travel time and fuel consumption, the various kinds of emissions can also be reduced significantly if traffic lights are optimized.

Finally, Figure 5.7 shows the traffic volume for the one-day simulation and during peak hours observed for the scenario studied in this work. Note that the main flows of agents go south-north-south rather than east-west-east, which reflects the demographics of the city.

5.7 Conclusion and Future Work

In this work, we analyzed the evolutionary optimization of traffic signals minimizing simultaneously travel time and fuel consumption on a large real-world scenario. We integrated a multi-objective evolutionary algorithm with the transport simulator MATSim and the emissions model simulator CMEM. We used as a case study the transport network of a 5×8 Km^2 area of Quito set with 70 signal lights, and simulated one day traffic of 20.000 agents moving according a two-leg mobility plan. We showed that there is a clear trade-off between travel time and fuel consumption when the signals are set with the same cycle length and are not coordinated (there is not offset between the start of the cycles). We also showed that the optimization of the signals allowing different cycle lengths between signals and coordinating them by properly setting their offsets can reduce significantly both travel time and fuel consumption. This reduces the range of the trade-offs between the two objectives. Further, we verified that the bi-objective optimization

(a) One day simulation.

(b) Peak hours (08h00 and 17h00).

FIGURE 5.7: Traffic Volume

approach produces better results than a single-objective approach that optimizes only travel time. We showed evidence that the single-objective algorithm is misled by the initially uncoordinated signals where larger cycle lengths allow shorter travel times, whereas combinations of coordinated signals with shorter cycle lengths lead to better travel times and lower fuel consumption. This was not an issue for the multi-objective optimizers because the second objective related to fuel consumption favors shorter cycle lengths even in uncoordinated signals.

As future works, we should improve the evolutionary algorithm to reduce its variance and enhance its reliability for short-term evolution and few fitness evaluations. Also, we should study other mobility plans and scenarios for the agents. Furthermore, in addition to optimizing traffic signals, we would like to add new variables and optimization criteria to study other important aspects of sustainable transport systems.

Chapter 6

Level of Service Optimization: Bi-Objective Formulation

The optimization of the transportation's level of service is addressed in this chapter. A scenario where a proportion of private and public transportation users move in the area of study was implemented and tested since a bi-objective perspective. A study of variables, trade-offs and emissions is also presented.

6.1 Introduction

Currently, around 54 percent of the world's population lives in urban areas. The world's population is expected to grow in the future and the proportion of the urban population as well. By 2050, the urban population is projected to increase to 66 percent (United Nations and Social Affairs, 2014). Population growth and urbanization trends have increased the demand for urban transportation and will continue to do so in the future. Often, existent traffic infrastructure and public transportation systems cannot cope with the increasing demand. Thus, dwellers tend to use their private means of transportation. Particularly, in emerging large cities, where the public transportation has not been able to satisfy the demand, the number of vehicles is increasing considerably and rendering mobility problems even worse (Gakenheimer, 1999). Congestion and delays add substantial costs that could be avoided, increases pollution and the risk of accidents.

The design optimization of urban transportation systems is a complex problem with multiple interrelated components that impose a continuous challenge for city planners. Evolutionary algorithms and other heuristics have proved effective to optimize difficult problems and are being investigated in the mobility and transportation domain. Some applications of single and multi-objective evolutionary algorithms to road network design, bus network optimization, and public transportation routing scheduling are as follow. In an early work(Bielli, Caramia, and Carotenuto, 2002), the authors use a single-objective genetic algorithm (GA) for bus network optimization. In (Hee, Kyung, and Hyuk, 2012) the authors solve a road network design problem (RNDP) using a bi-level optimization approach that reflects the different objectives between planners and network users. They focused on the design of a small network with six links and six nodes using a multi-objective

GA and optimizing three objectives related to travel time, fuel consumption, and accessibility to network's nodes. Other works focus on public transportation routing and scheduling. In (Arbex and Cunha, 2015) the authors use a bi-objective approach to minimize both passengers and operators costs on a benchmark network (15 nodes and 21 links). In (Ma et al., 2017) the authors studied stop selection, line planing and timetables for customized bus (CB).They used a single-objective GA to solve a weighted fitness function considering passenger, operator, and societal perspectives and test their model on real network areas that cover $3km^2$ and $18km^2$. Other works focusing on frequency setting and vehicle scheduling problem employee genetic algorithms and a bi-level approach. Regarding frequency setting problem, in (Yu, Yang, and Yao, 2010), at the upper level, the frequency routes are optimized minimizing the passenger's travel time meanwhile in the lower level an assignation of transit trips to bus optimal route strategy is set. A single objective genetic algorithm (GA) is implemented to search the frequency of each route. The results obtained show that the optimal solution can decrease total travel time by about 6% compared with the current situation. They used a heuristic where the optimal strategy and the expected total travel time from each node to the destination node are computed, and the demand is assigned to the network according to the optimal strategy from all origins to the destination. In another study (Verbas and Mahmassani, 2015), the authors propose a bi-level model where at the upper level a frequency setting problem is addressed minimizing waiting times and use a heuristic model running a simulation at the lower level according to a demand and frequency. Both previous works used one pre-defined demand, a macro-model and simulation schema to study the effects of a frequency setting in travel time and waiting time respectively. Regarding scheduling problems, in (Farhan Ahmad, 2005), the authors state a bi-level approach to solve a bus scheduling problem for overlapping routes. In the first level, the model minimizes the number of buses required for each route individually and considering load feasibility. In the second level, given the fleet size computed in the first level, the fleet size is minimized again by considering all routes together by using a single objective genetic algorithm (GA). They tested the model in a real-world network with a total of 60 nodes and 70 links. Link distance in time units and symmetric demand matrix gives the evaluation of each solution. The results showed that the reduction in fleet size is not significant as it was expected during model formulation stage due to little overlapping of routes in the test network.

In this work, we study the urban transportation system under various proportions of private and public transportation users, aiming to understand the conditions to achieve different levels of service and their implications on the optimal configuration of the public transportation, fuel consumption (energy), and the impact on the environment.

The level of service (LoS) is a quality measure describing operational conditions within a traffic stream, generally in terms of service measures such as speed and travel time, freedom of maneuver, traffic interruptions, and comfort and convenience (Board, 2000). In this work, we define LoS in terms of

traffic density, which is determined by the total number of vehicles that circulate in the city given by the proportions of the population that use public and private transportation.

To make public transportation more attractive, city planners need to guarantee a reliable service with appropriate frequency and capacity of buses to satisfy the mobility demand. For a given proportion of public transportation users, we should find a proper configuration of bus capacities and departure times between buses. Different proportion and configurations will impact travel time, fuel consumption, costs and emissions, in addition to traffic density. Most of these criteria are in conflict with each other. For example, the transportation service administration would like to maximize the use of the infrastructure minimizing costs and energy. Increasing the frequency of buses can increase the capacity of the system. However, this could carry additional costs. Users would like to minimize their travel times. This could be an incentive to use private transportation but could increase their costs. In addition, people living in zones where there are public transportation routes would like to minimize emissions or pollution. A way to improve the flow is reducing the volume of traffic or density. However, reducing density has impacts in the total travel time of the users because they need to use public transportation. These trade-offs suggest that compromised solutions can be found using a multi-objective optimization approach.

We use a multi-objective evolutionary algorithm to search combinations of number of private/public transportation users, capacity of buses, and time interval between bus departures, minimizing traffic density and travel time simultaneously. In order to have a broad vision of the possible solutions to the problem and their level of service, at this time we assume that buses are always available in different capacities, do not constraint their total number, and do not consider cost.

We conduct our study using a scenario based on a real traffic network from Quito city (Ecuador). Quito's population has grown considerably in the last years. Currently, its urban transportation system is highly congested due to an increase of private vehicles combined with a poorly satisfied demand by public transportation. It is important to find alternative configurations of the urban transportation system to improve the level of service and quality of life in the city. In this work, we model the mobility of 27.000 agents that use private and public transportation on an area of $40 km^2$, and simulate their mobility using MATSim, a multi-agent transport simulator.

6.2 Method

6.2.1 Components & Overview

We follow the design optimization and method approach described in section 2.2 and illustrated in Fig 2.1.

6.2.2 Problem Definition

Given a total fixed number of travellers that can use either public or private transportation, a number of bus lines and a set of schedule periods, the problem consists in finding optimal configurations of the public transportation system in terms of bus capacities and headways (time between bus departures in the same line) for a range of proportions of public transportation users minimizing simultaneously travel time and traffic density.

The problem is formulated as follows

$$
\begin{aligned}
&\text{minimize} && f(x) = (f_1(x), f_2(x)) \\
&\text{subject to} && \\
& && x = (n_{Pc}, n_{Pt}, c, h) \\
& && n_A = n_{Pc} + n_{Pt} \\
& && 0.2n_A < n_{Pt} < 0.8n_A \\
& && c = \{c_i^j \in C\}, i = 1, \cdots, n_B \text{ and } j = 1, \cdots, n_S \\
& && h = \{h_i^j \in H\}, i = 1, \cdots, n_B \text{ and } j = 1, \cdots, n_S
\end{aligned}
$$

Function $f_1(x)$ measures travel time of all travellers and function $f_2(x)$ measures the traffic density, x is the vector of decision variables, n_{Pc} is number of private transportation travellers, n_{Pt} is the number of public transportation users, c_i^j is the capacity (maximum number of passengers) of the buses assigned to line i at schedule period j, n_B is the number of bus lines, n_S is the number of schedule periods, c is the vector of capacities associated to all bus lines and schedule periods, h_i^j is the headway departure or time between departure of buses on line i for schedule period j, h is the vector of headways associated to all bus lines and schedule periods, C is a finite set of bus capacities, H is a finite set of headways expressed in minutes.

6.2.3 Area of Study

We study the conditions to achieve different levels of service in Quito city (Ecuador). The geographical area of study includes the central zone of the city, where several universities, big malls, two parks, a stadium, and the main hub of the private and public transportation infrastructure are located. The area of study also includes residential zones and the business district. The total area represents approximately $40 km^2$. Fig. 6.1 shows the extension of the geographical area.

The population in Quito was around 2 million in 2010 (INEC, 2010b), of which 80% move every day to accomplish various activities. Main mobility reasons according to (DMQ, 2012) are 31% for work, 32.5% for study (all educational levels) and 36.5% for other different activities. It is estimated that 47% of traffic flows go to the zone we study in this work (GrupoFaro, 2010). From this percentage, 74.5% use public transportation and 25.5% private cars. In our study, we simulate a sample of 27.000 agents.

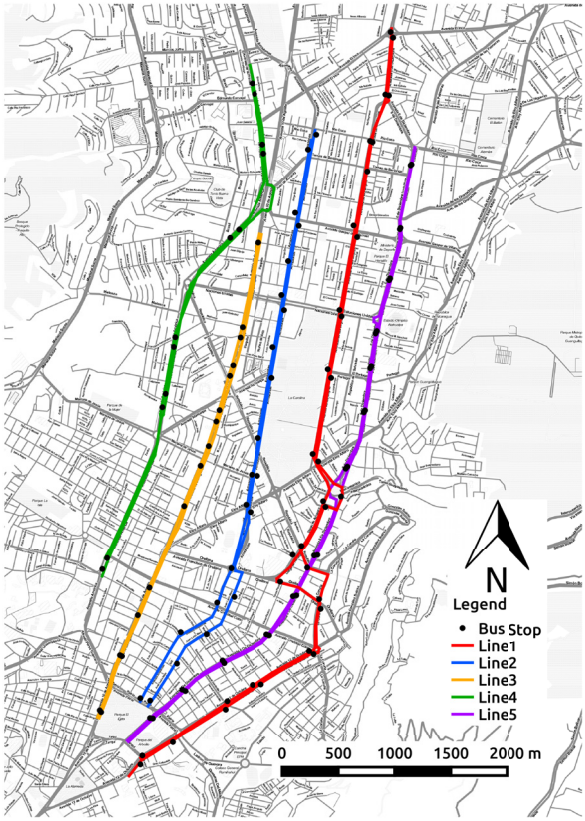


FIGURE 6.1: DMQ BRT Scenario

The massive passenger transportation in Quito is performed by private and public companies. The Passenger Transportation Metropolitan Company is a public company in charge of the Bus Rapid Transit (BRT) service. In 2015, approximately 650.000 passengers were transported daily by the different BRT lines (DMQ, 2015). In our model, we consider five main BRT corridors (see Fig. 6.1) which are the most demanded and congested routes.

6.2.4 Simulation and MOEA Integration

MATSim performs a micro-simulation of agents that move on a transportation system producing information about agent routes and movements. MATSim requires as inputs the model of the transport network infrastructure, public transportation configuration, and the agents' mobility plans.

Fig. 6.2 shows the integration between the simulator and the evolutionary algorithm (EA). The EA searches optimal settings for capacity and headway bus line configuration for a given proportion of agents that will use public transportation.

To evaluate a solution, MATSim sets the bus line capacities and headways with the solution provided by the optimizer, computes the routes of the agents based on heuristics, and simulates the mobility of all agents (private and public transportation) following their mobility plans. MATSim simulates the public transportation by setting stop locations, lines, schedule,

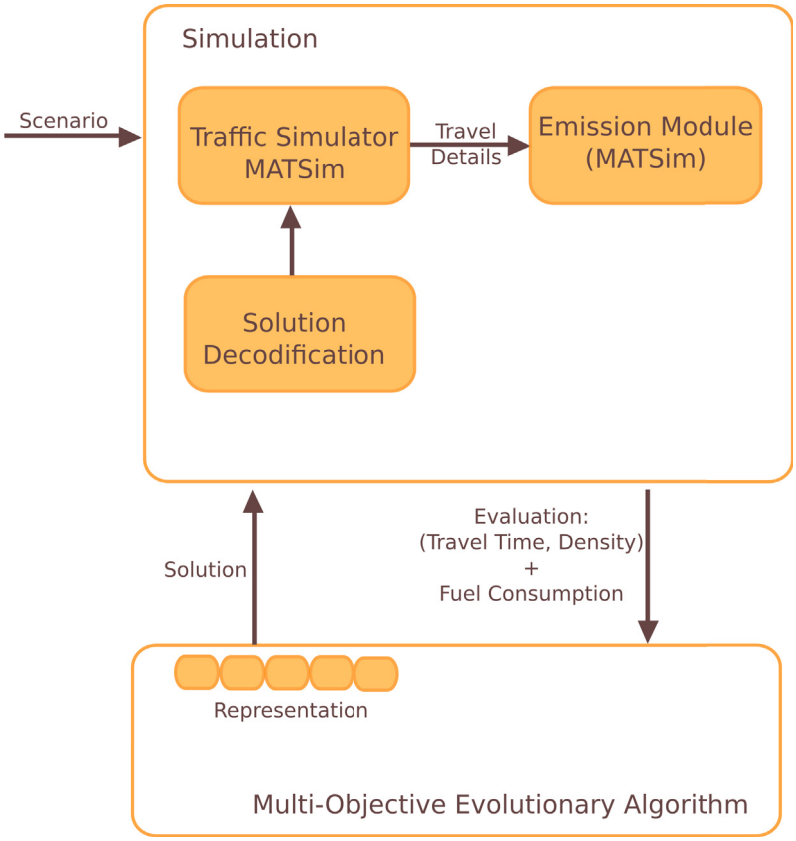


FIGURE 6.2: Simulation and EA integration

routes, departures, and vehicles. A line consists of two or more routes defined over the network links and a list of departure times with a reference to a particular bus specification (for example, capacity, length, technology) (Horni, Nagel, and Axhausen, 2016). MATSim emission module (Hülsmann et al., 2011) computes the fuel consumption and emissions for cars and buses by connecting MATSim output to comprehensive database Handbook on Emission Factors for Road Transportation (HBEFA) (Keller and Wuthrich, 2014). The output collected from the simulator is used to calculate travel time and traffic density, which are passed back to the optimizer as objective values to compute the fitness of the solution.

6.3 Evolutionary Algorithm

In this work, we use the Adaptive ε-Sampling and v-Hood (AεSvH) algorithm to search optimal solutions. AεSvH algorithm is described in section 2.5.4.

6.3.1 Solution Representation

In this work we consider private cars and city buses as the only means of private and public transportation, respectively. The main components of our representation are the number of travellers that use public transportation n_{Pt}

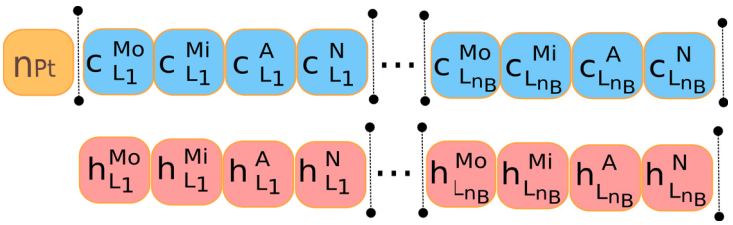


FIGURE 6.3: Solution Representation

and the bus lines configurations given by capacities **c** and headways **h**. In this work n_{Pt} can take values in the range $[0.2n_A, 0.8n_A]$ in intervals of $0.1n_A$, where n_A is the total number of travellers or agents. The number of private transportation travellers is $n_{Pc} = n_A - n_{Pt}$. The n_B bus lines are identified by the set of indexes $L = \{L_1, \cdots, L_{n_B}\}$ and the n_S schedule periods by the set of indexes $S = \{M_o, M_i, A, N\}$ for morning, midday, afternoon, and night. A solution for the evolutionary algorithm is represented by the following integer variables

$$(n_{Pt}, c, h)\,, \tag{6.1}$$

where

$$c = \{c_i^j \in C\}, i = 1, \cdots, n_B \text{ and } j = 1, \cdots, n_S \tag{6.2}$$
$$c = \{c_{L_1}^{M_o}, \cdots, c_{L_1}^{N}, \cdots, c_{L_{n_B}}^{M_o}, \cdots, c_{L_{n_B}}^{N}\}$$

$$h = \{h_i^j \in H\}, i = 1, \cdots, n_B \text{ and } j = 1, \cdots, n_S \tag{6.3}$$
$$h = \{h_{L_1}^{M_o}, \cdots, h_{L_1}^{N}, \cdots, h_{L_{n_B}}^{M_o}, \cdots, h_{L_{n_B}}^{N}\}$$

Here, c_i^j is the capacity of bus (number of passengers) of line i for schedule period j, $c_i^j \in C = \{0, 1, 2\}$ where 0, 1 and 2 denote small, medium and large capacity, respectively. h_i^j is the headway departure or time between departure of buses on line i for schedule period j. $h_i^j \in H = \{0, 1, \cdots, 11\}$, to represent headways from 5 to 60 minutes in steps of 5. Fig. 6.3 illustrates the representation of solutions used by the evolutionary algorithm.

6.3.2 Operators

To create offspring we follow the representation described above and apply uniform crossover (UX) with probability P_c, using a mixing ratio pCv. For mutation we first decide whether to mutate n_{Pt} with probability $P_m^{n_{Pt}}$ or capacities and headways (**c**,**h**) with probability $1 - P_m^{n_{Pt}}$. When capacities and headways are mutated, we traverse capacities and headways for all lines and schedule periods and mutate them with probability P_m per variable in (**c**,**h**). The mutation operators increase or decrease randomly with equal probability the code of the variable in which operates.

6.4 Evaluation Functions

The fitness functions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ for travel time and traffic density referred in the problem definition of section 6.2.2 are computed from the outcome of the simulation. In the following we express them in terms of the data generated from the simulator.

6.4.1 Travel Time f_1

The travel time *tt* is given by the total travel time of private and public transportation and expressed by

$$tt = tt_{pc} + tt_{pt} \tag{6.4}$$

The average travel time for private (cars) is expressed by

$$tt_{pc} = \frac{1}{n_{Pc}} \sum_{i=1}^{n_{Pc}} \sum_{l=1}^{L} t_{il}, \tag{6.5}$$

where t_{il} represents the travel time on link l for vehicle i and L is the total links of the route.

The average travel time for public transit users is computed by

$$tt_{pt} = \frac{1}{n_{Pt}} \sum_{i=1}^{n_{Pt}} t_i^a - t_i^d, \tag{6.6}$$

where where t_i^a represents the arrival time to its destination and t_i^d is the departure time from the origin of the trip for an agent *i*. This time includes walking time to the bus stop, waiting time, and bus riding time.

6.4.2 Density f_2

One measure of the effectiveness to define levels of service is density. Density describes the proximity of vehicles to each other, which is the principal influence on freedom to manoeuvre. Thus, it is an appropriate descriptor of service quality(Roess, Prassas, and McShane, 2011). Density is a relationship given by $D = v \,/\, S$ where v is the volume or flow rate, S is the average passenger-car speed(km/h) and D is the density (pc/km/ln)(Roess, Prassas, and McShane, 2011). For our scenario, density is computed by

$$density = \frac{N_{TripsPc} + N_{TripsPt}}{AvgSpeed} \times \frac{1}{AvgNLinks}, \tag{6.7}$$

where $N_{TripsPc}$ and $N_{TripsPt}$ are the total number of trips of private vehicles and public transportation, respectively, computed on the routes taken by agents. *AvgSpeed* is the total average speed of vehicles during the scenario duration time. *AvgNLinks* is the average of links traversed for each agent during its trip.

6.5 Simulation Results and Discussion

6.5.1 Experimental Setup

We optimize 5 BRT lines, splitting the schedule of buses in four time periods from 06h00 to 09h00, 09h00 to 16h00, 16h00 to 20h00, and 20h00 to 22h00. These correspond to the schedule periods set $S = \{M_o, M_i, A, N\}$ for morning, midday, afternoon, and night. For each schedule period, the algorithm searches appropriate capacities and headways (frequencies) of buses. Thus, the number of variables is 41, one for the proportion of public transportation users n_{Pt}, **5** lines $\times$ **4** schedule periods for capacities **c** and **5** lines $\times$ **4** schedule periods for headways **h**. Capacity categories for BRT vehicles are small, medium and large with 25, 70 and 115 passengers respectively. Headway is split into 12 categories from 5 to 60 minutes in steps of five minutes.

For the evolutionary algorithm, we set population size to 100 and initialize it randomly. For the operators, we set crossover rate to Pc = 1.0 with mixing ratio probability pCv=0.3. This means that all individuals undergo uniform crossover and offspring keeps in average 70% of the variables of one parent and 30% of the other. Mutation rate $P_m^{n_{Pt}}$ is set to 0.1 to mutate n_{Pt} and $\mathbf{1} - P_m^{n_{Pt}} = \mathbf{0.9}$ to mutate variables in (**c**,**h**). Mutation rate per variable in (**c**,**h**) is Pm = 1/n, where $n = \mathbf{40}$ is the number of variables in (**c**,**h**). Thus, in average, mutation changes the variable n_{Pt} corresponding to the number of public transportation users for 10% of the population and the variables corresponding to capacity or headways of one of the bus lines for 90% of the population. We conduct ten runs of the evolutionary algorithm fixing the number of generations to 100. We perform the experiments on a 64 bits Rocks cluster with 12 nodes, of which 9 are six core 2GHz and 3 are sixteen core 2.8GHz. The population is evaluated in parallel and the evaluation of one individual takes in average 14 minutes.

To evaluate each solution, we run Matsim simulating the mobility of $n_A = \mathbf{27.000}$ agents. The mobility plan for each agent consists of two main trips or legs: 1) from home to work, study, or others and 2) from work, study, or others to home. The plans are designed so that all agents move first from each home location to different points along the area of study. Those points are facility locations such as universities, workplaces and others like malls, and parks. In their second trip, the agents move back home. The distribution of home locations, workplaces, and education facilities for the mobility plan were chosen taking into account census data and a previous mobility study(Demoraes, 2005). We prepare in advance the settings of public transportation configuration: lines, routes, and stop locations. The public transportation configuration consists of five BRT lines, 112 stop locations and two routes per each bus line. When Matsim recieves the solution proposed by the evolutionary algorithm, i.e. a vector of values corresponding to variables (n_{Pt}, **c**,**h**), the agents that will use public transportation are choosen randomly with probability n_{Pt}/n_A. The other agents are set to use private cars. Also, the capacities and headways for the five lines are set accordingly to the values suggested by the evolutionary algorithm.

TABLE 6.1: **Automovile distribution (fuel=gasoline and weight $\leq$ 2Tons)**

Year-Category	%	Year-Category	%
2000-2002 | Euro2	24.40	2008-2010 | Euro5	13.70
2002-2004 | Euro3	7.80	2010-2012 | Euro6	23.80
2004-2006 | Euro4	12.20	2012-2014 | Euro6	5.20
2006-2008 | Euro4	13.00		

To compute emissions, we have selected seven categories of vehicles for private and one for public transportation. Table 6.1 shows the distribution of vehicle categories for private transportation chosen for the scenario, which is in accordance with census transportation data(INEC, 2010b) for Quito city. We assign a category of vehicle to each agent randomly following this distribution. For public transportation, we have chosen a *heavy goods vehicle* HGV type which weight goes from 14 to 20 tons and diesel EuroII as fuel/technology type. Vehicle characteristics enable to identify the emission factors from HBEFA database. To compute warm emissions, MATSim combines the vehicle information with the kinematic features such as speed and stop duration obtained from the simulation.

6.5.2 Comparison between AεSεH and AεSvH

In this section we compare the proposed algorithm A$\boldsymbol{\varepsilon S v H}$ that creates neighborhoods in decision space to perform recombination with the A$\boldsymbol{\varepsilon S \varepsilon H}$ that creates neighborhoods in objective space.

First, we compute the hypervolume of the non-dominated set of solutions at each generation to verify their evolution. The reference point to calculate the hypervolume is fixed to the worst values in each objective, obtained from the Pareto set of non-dominated solutions found by the algorithms in all runs. Fig. 6.4 shows box plots of the hypervolume over the generations for the ten runs of both algorithms. In blue we show results by A$\boldsymbol{\varepsilon S \varepsilon H}$ and in red results by AεSvH. Note that the hypervolume rapidly increases with the generations, showing that evolution is proceeding as expected. However, the hypervolume by AεSvH is significantly larger than A$\boldsymbol{\varepsilon S \varepsilon H}$. This means that A$\varepsilon$S$v$H can find sets of non-dominated solutions with better properties of convergence and diversity than A$\boldsymbol{\varepsilon S \varepsilon H}$.

Figure 6.5 shows the hypervolume of a typical run by both algorithms starting them with the same random seed. Figure 6.6 shows the sets of non-dominated solutions found by the algorithms for the same runs. Solutions by A$\boldsymbol{\varepsilon S \varepsilon H}$ are shown in black, whereas solutions by AεSvH are colored according to $\boldsymbol{n_{Pt}}$ values. Note that AεSvH can find solutions for all proportions of public transportation users $\boldsymbol{n_{Pt}}$, whereas A$\boldsymbol{\varepsilon S \varepsilon H}$ loses diversity in variable space and is not able to find solutions for large values of $\boldsymbol{n_{Pt}}$, particularly for **0.7** and **0.8**. In addition, note that for all values of $\boldsymbol{n_{Pt}}$ solutions by AεSvH converge to better values in objective space than solutions by A$\boldsymbol{\varepsilon S \varepsilon H}$.

The two set coverage index is a metric used to compare non-dominated sets found by muti-objective evolutionary algorithms and expressed in Eq.6.8.

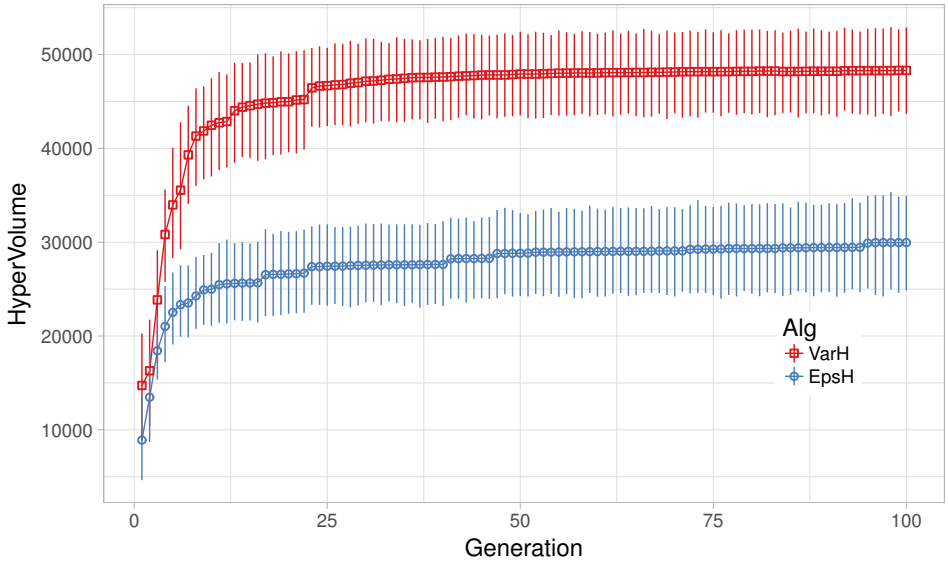


FIGURE 6.4: Hypervolume Over Generations

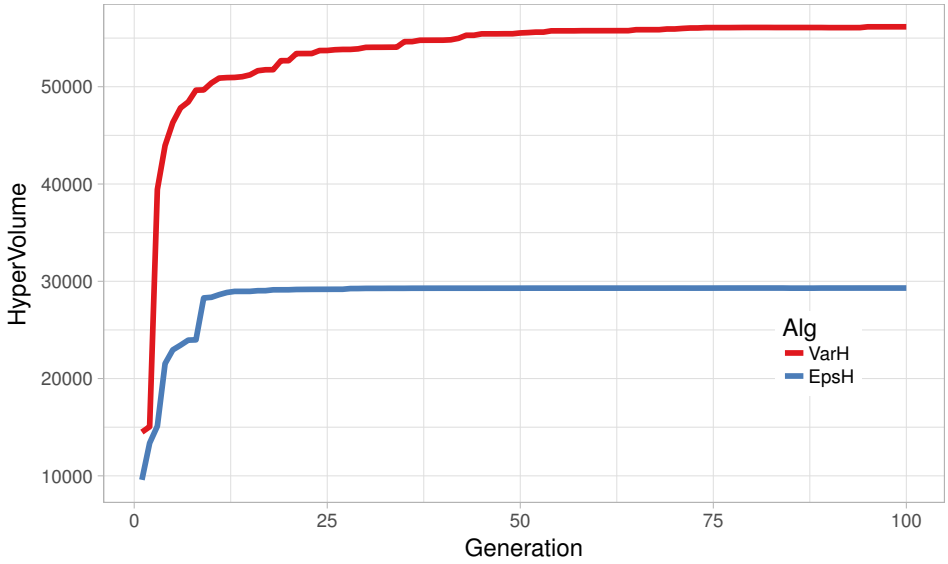


FIGURE 6.5: Hypervolume Over Generations of One Run.

Given two Pareto non-dominated sets, A and B, set-coverage C(A,B) counts the number of solutions in the set B that are dominated by the solutions in the set A upon the number of solutions in the set B.

$$C(A,B) = \frac{|\{b \in B \mid \exists a \in A \ : \ a \preccurlyeq b\}|}{|B|} \tag{6.8}$$

Table 6.2 shows the average set-coverage values of the Pareto non-dominated solutions found by both algorithms. Note that in average solutions by AεSvH dominate many more solutions by A$\varepsilon S\varepsilon H$, which is in accordance with the hypervolume values reported above.

In the following we focus our analysis on solutions found by AεSvH.

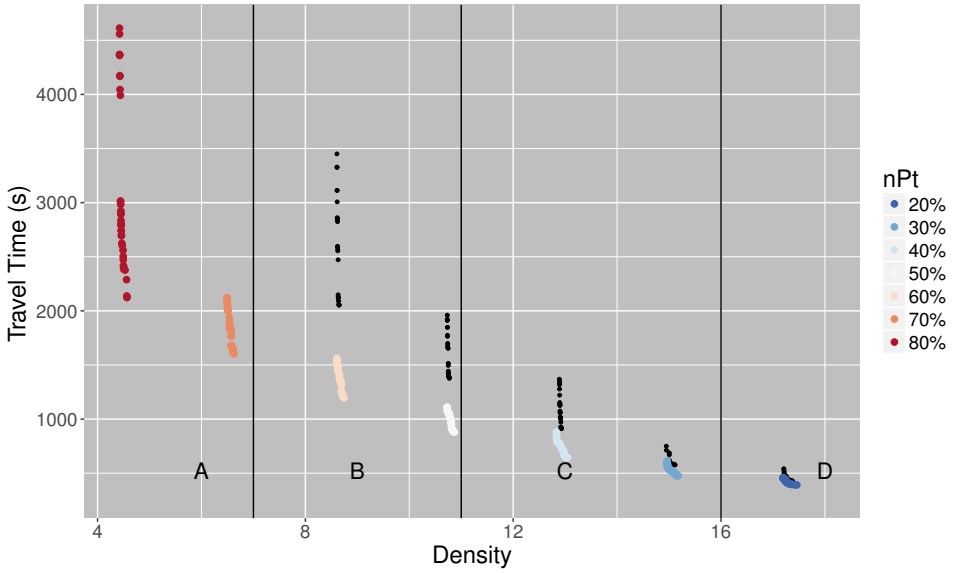


FIGURE 6.6: Pareto Optimal Set of one run.

TABLE 6.2: Two Set Coverage Index (C)

	mean	sd
C(AεSεH,AεSvH)	0.24	0.18
C(AεSvH,AεSεH)	0.50	0.31

6.5.3 Trade-offs

Next, we analyze trade-offs between objectives. Unless stated otherwise, we use in our analysis the set of non-dominated solutions found in the ten runs of the algorithm. The correlation values between objectives travel time (tt) and density (dens) are shown in Table 6.3, where negative correlations suggest trade-offs between objectives. This table also includes the correlation between the objectives and fuel comsuption(fc), which is measured but not used for optimization.

TABLE 6.3: Objective Correlation Matrix

	tt	dens	fc
tt	—	-0.91	-0.92
dens	-0.91	—	0.99
fc	-0.92	0.99	—

Fig. 6.7 shows the trade-off between *travel time* and *density*, which is in agreement with the negative correlation shown in Table 6.3. This figure is colored according to the proportion of agents n_{Pt} that use public transportation. Also, it includes vertical lines to mark ranges of the levels of service (LoS) computed from the traffic density. To compute LoS we use the six levels of service defined by The Highway Capacity Manual (Board, 2000), designated by the letters A through F, with A being the highest level of service and F the lowest. Table 6.4 shows the different levels according to density. In the figure, the lines correspond to ranges of LoS A-D.

TABLE 6.4: Levels of service for basic freeways segments

Levels of Service	Density (pc/km/ln)
A	0-7
B	7-11
C	11-16
D	16-22
E	22-28
F	>28

From this figure it can be seen that high densities of traffic correspond to smaller proportions of public transportation users n_{Pt} and vice versa, as expected. Regarding levels of service, note that the best level of service A requires that 70% and 80% of the agents use public transportation, level of service B can be achieved with 50% and 60% of these agents, level C with 30% and 40%, and the worst level D requires 20% when more people uses private cars that leads to the highest density of traffic.

From the same figure, note that solutions for a given value of n_{Pt} are clustered in density and travel time. An analysis of these clusters and their trade-off ranges provide valuable information. To illustrate, note that within each cluster there is no much difference in density, but travel time can change significantly, especially for larger fractions of public transportation users. The difference in travel time for each value of n_{Pt} is due to the various configurations of bus headways and bus capacities. Within each cluster, solutions with lower travel time are achieved by scheduling large capacity buses with high frequency. As mentioned above, larger trade-off ranges in travel time are seen for larger n_{Pt}, where public transportation is used more. For example, the trade-off range in average travel time is around 500 seconds for cluster $n_{Pt} = \mathbf{70\%}$ and 150 seconds for cluster $n_{Pt} = \mathbf{30\%}$. This clearly shows ranges of improvement in travel time that can be achieved by optimizing public transportation with almost no impact to traffic density for a given value of n_{Pt}. On the other hand, notable changes in density are observed between clusters, for different values of n_{Pt}. This suggests that to reduce density a significant number of users should be encouraged to use public transportation instead of private cars.

In addition to travel time and density, we also estimate fuel consumption though at this time we do not use it as an objective value for optimization. Fig. 6.8 shows *travel time* and *fuel consumption*, coloring solutions by the LoS. Note that there is a perceptible trade-off between fuel consumption and travel time, in agreement with the negative correlation shown in Table 6.3. Looking at the levels of service, note that in general solutions with the best LoS, such as A and B, have lower values of fuel consumption, but higher travel times. This is because high levels of service imply low traffic densities, and vice versa. High levels of service are achieved under scenarios with larger fractions of public transportation users, where, on average takes the user longer to get to their destinations, but less private vehicles circulate the

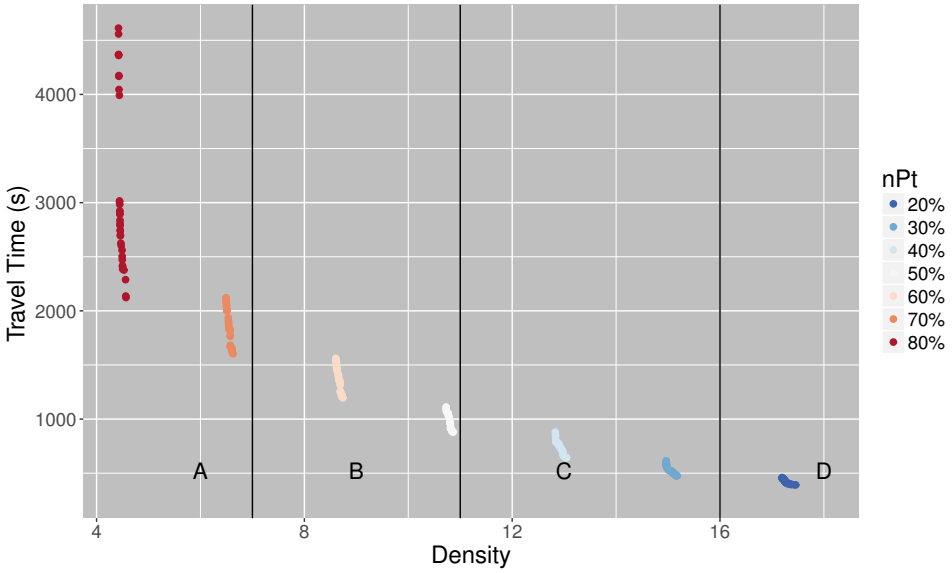


FIGURE 6.7: Travel time and density, colored by fraction of public transport users n_{Pt}. Black dots show POS founded by the original algorithm.

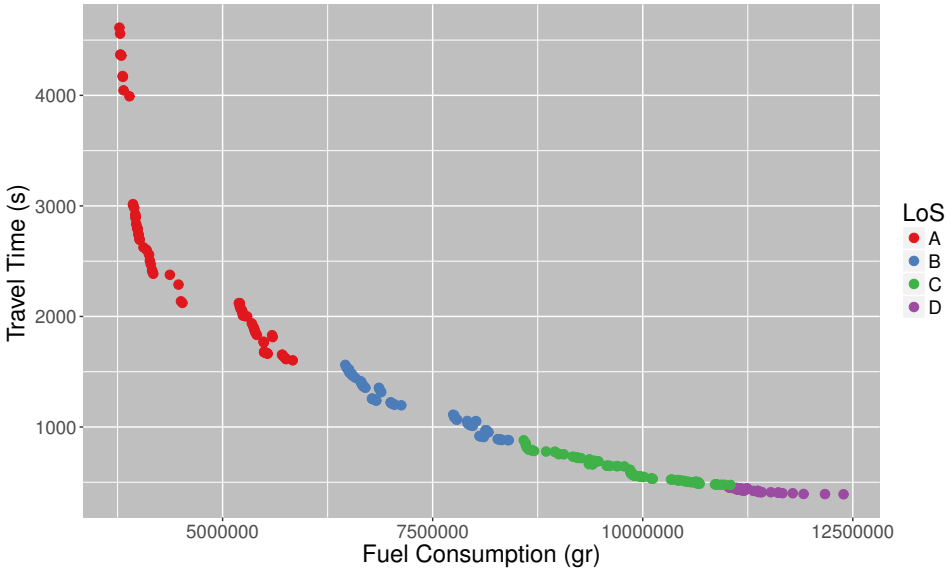


FIGURE 6.8: Fuel consumption and travel time, colored by level of service LoS.

city. On the other hand, solutions with worst levels of service, C and D, in general show better travel time but worse fuel consumption. Thus, the increase in fuel consumption is because more people uses their cars instead of public transportation.

Due to the high positive correlation between density and fuel consumption, shown in Table 6.3, a clear correspondence between Fig. 6.7 and Fig. 6.8 emerges. Looking at these two figures, note that for a given fraction n_{Pt} of public transportation users the difference in fuel consumption can become significant although density does not change much. This can be seen clearer

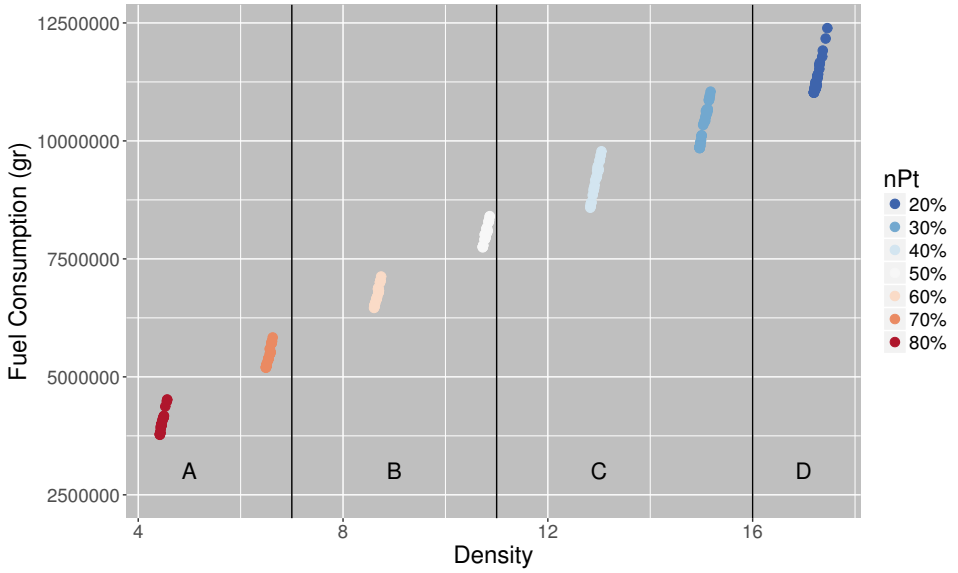


FIGURE 6.9: Fuel consumption and density, colored by fraction of public transport users n_{Pt}.

in Fig. 6.9 that shows *fuel consumption* and *density*, coloring solutions according to n_{Pt}. The strong positive correlation between density and fuel consumption suggests that the latter could be used instead of the former as objective value for optimization.

6.5.4 Emissions

Another important aspect of sustainable public transportation is its environmental impact. As stated above, larger fractions of public transportation users n_{Pt} reduce overall fuel consumption in the city as shown in Fig. 6.9. Also, best levels of service and low levels of fuel consumption can be achieved simultaneously, as shown in Fig. 6.8.

In addition to fuel consumption, we collect data from MATSim's emissions module. In this work, we analyze particulate matter (PM) emissions. The concentration of PM is composed of engine exhaust gas emissions as well as abrasion and suspension. Regarding health effects, the smaller the particle, the worse the impact on humans. The effects range from damages of the respiratory system to carcinogenic effects.

Fig. 6.10 shows travel time against PM emissions of the non-dominated solutions found by the evolutionary algorithm, coloring by n_{Pt}. Note that there is a trade-off between travel time and PM emissions within each cluster n_{Pt}. For a given fraction of n_{Pt} the number of private cars is the same and the range of PM emissions observed is determined by the different configurations of headways (frequency) of buses. Thus, for a given n_{Pt}, increasing the frequency of buses reduces travel time but increases PM. The difference in travel time is more notorious for large n_{Pt}, where most people use public transportation. In this work we model buses with diesel as fuel, which reflects the current technology for most buses in Quito. This results suggest that a change of technology for buses, for example from diesel to electric,

is essential to achieve reasonable travel time and low traffic densities with minimum PM emissions.

From Fig. 6.10 also note that the upper bound of PM increases with smaller fractions of n_{Pt}. This is due to configurations with high frequency of buses combined with a large number of private cars. In fact, the solution with highest PM emissions for $n_{Pt} = \mathbf{80\%}$ includes a configuration where buses are run with slightly higher frequency than the solution with highest PM emissions for $n_{Pt} = \mathbf{30\%}$. So, the difference in PM emissions between these two extreme solution is mostly due to the difference in the number of private cars. High frequency configurations of buses for low demand scenarios slightly reduce travel time and are likely to have a lower occupancy rate. These solutions appear as non-dominated because at this time we do not consider cost nor take into account occupancy rate of the buses.

Fig. 6.11 and Fig. 6.12 show box plots of the number of public transportation trips and PM emissions, respectively, grouped by n_{Pt}. From these figures, note that the median of number of trips and PM emissions increase with n_{Pt} from **20%** to **50%**, and are similar for fractions **60%** and above.

We analyze with more detail the best and second best solutions in travel time found by the algorithm for $n_{Pt} = \mathbf{80\%}$, denoted as A8b and A8a as shown on Fig. 6.10. Note that there is a considerable difference in PM emissions between them. Fig. 6.13 shows separately the configurations of bus capacities (top) and headways (bottom) of all five lines and schedules for these solutions. Capacities and headways are colored according to their code. Together to the identifier of the solution we also include the value of PM emissions computed for public transportation, which are lower than those reported in Fig. 6.10 where the value of PM emissions include private cars as well. Note that overall the frequency of buses is higher in A8b than in A8a, which is reflected in the values of PM emissions. In addition, note that to satisfy the high demand of public transportation these solutions require low capacities and relatively low frequencies for some lines and schedules, see for example Line3 in the afternoon schedule. This illustrates the relevance of assigning buses of different capacities and frequencies to different schedules.

The geo-located differences in PM emissions between A8b and A8a are illustrated in Fig. 6.14, where red color indicates an increase of PM emissions in A8b and blue a reduction. Note, for example, that the levels of PM increase considerably on the network segments where bus lines 4 and 3 are located (see Fig. 6.1 for the location of bus lines), where A8b has higher frequency than A8a as shown in Fig. 6.13.

6.6 Conclusions and Future Work

This work studied levels of service in urban transportation on a scenario of a real world traffic network. We coupled a multi-objective evolutionary algorithm with the multi-agent transport simulator MATSim to perform a bi-level search of the proportions of private/public transportation users and optimal configurations of buses capacities and headways associated to each proportion. The multi-objective evolutionary algorithm minimized travel time and

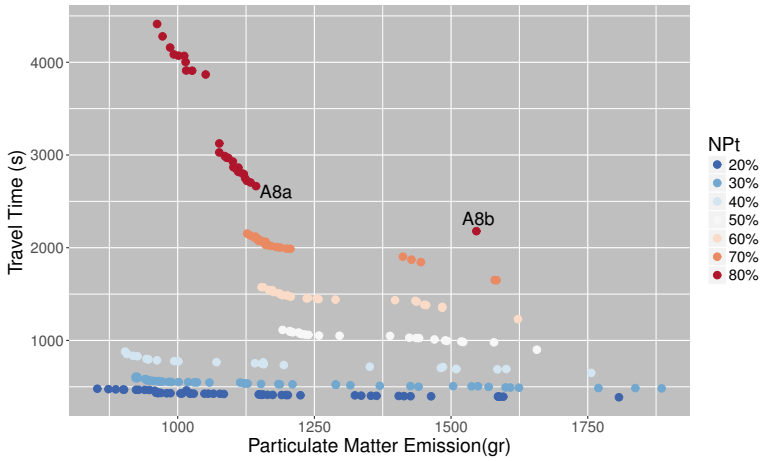


FIGURE 6.10: Trade-off TT and PM by n_{Pt}

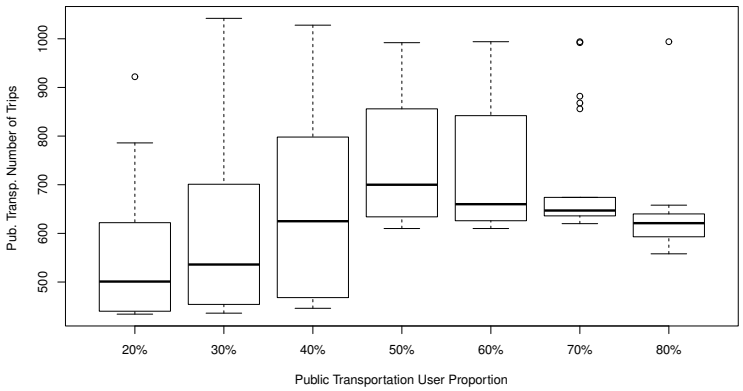


FIGURE 6.11: Number of Public Transportation Trips by n_{Pt}

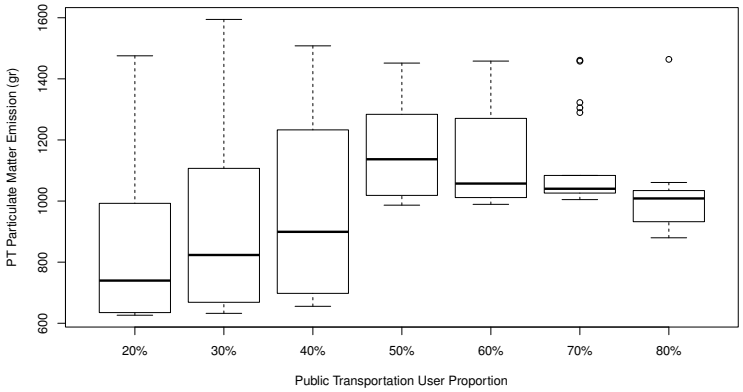


FIGURE 6.12: Public Transportation PM Emission by n_{Pt}

density simultaneously. We defined levels of service in terms of density and analyzed it in conjunction with the trade-offs between objectives on the set of non-dominated solutions found in ten runs of the algorithm. We also analyzed fuel consumption and the effects of particulate matter (PM) emission

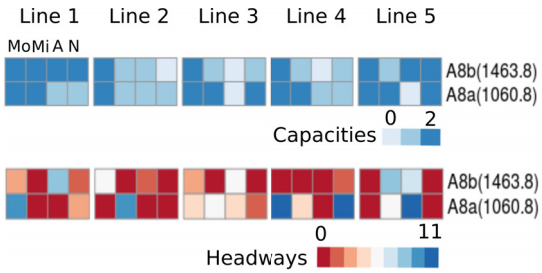


FIGURE 6.13: Configuration of bus Capacities (top) and Headways (bottom) of solutions A8b and A8a. Schedules: Mo, Mi, A, N (morning, midday, afternoon, night). Capacities: 0, 1, 2 (small, medium, large). Headways: 0, $\cdots$, 11 (5 min,$\cdots$, 60 min).

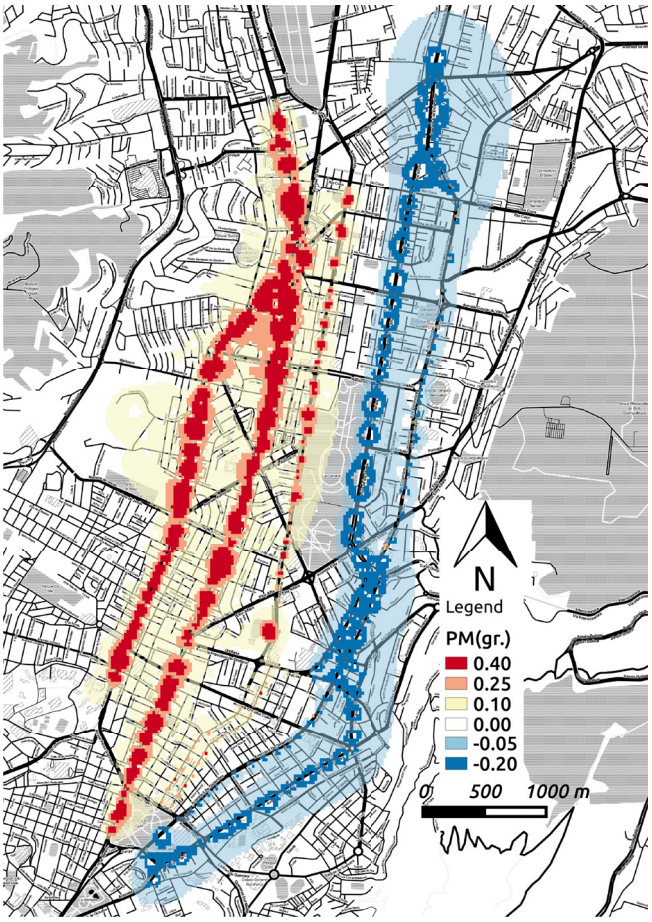


FIGURE 6.14: PM Emission Difference Between A8b and A8a

over the area. The results of this study provide valuable insights to understand better the conditions that can lead to different levels of service in the city. It also suggests ways to improve mobility combining private and public transportation. Our results, clearly show that to improve LoS, more travelers must use public transportation, but better LoS do not necessary implies a reduction in PM because it depends on the BRT headways and the diesel engines of the buses. Results also show that the trade-off between

travel time and fuel consumption agree with several configurations in capacities and headways according to the proportion of users. Also, PM emission shows that some of that settings can be environmentally friendly and user convenient. That information is worthy for decision makers due to the compromise between that objectives. In addition, results suggest that a change of technology for buses is essential to achieve reasonable travel time and low traffic densities with minimum PM emissions.

In the future, we would like to refine the model to consider constraints in resources. Also include other objectives such as cost, waiting time, bus occupancy rates, and evaluate the impact of a transition towards hybrid and electrical cars for private and public transportation. We would also like to extend the definition of service to cover other criteria besides density.

Chapter 7

Level of Service: Three-Objective Formulation

The optimization of the transportation's level of service since a three-objective approach is studied in this chapter. Here the objectives to minimize are travel time, traffic density and fuel consumption. Similarly, a study of variables, trade-offs and emissions is also presented.

7.1 Introduction

For this work, we use the same scenario employed for the bi-objective optimization problem discussed in the previous chapter. Here our interest is to study fuel consumption as a new optimization objective, analyze the trade-off if any with density and travel time and the scalability effects in the scenario. Also, we intend to study and understand the impact of Bus Rapid Transit (BRT) line headways considering the minimization of density, travel time and fuel consumption simultaneously. The method, evolutionary algorithm, representation and operators are the same as used by bi-objective formulation.

7.2 Evaluation Functions

7.2.1 Travel Time

The travel time *tt* is given by the total travel time of private and public transportation and expressed by

$$tt = tt_{pc} + tt_{pt} \tag{7.1}$$

The average travel time for private (cars) is expressed by

$$tt_{pc} = \frac{1}{N_{Pc}} \sum_{i=1}^{N_{Pc}} \sum_{l=1}^{L} t_{il}, \tag{7.2}$$

where t_{il} represents the travel time on link l for vehicle i and L is the total links of the route.

The average travel time for public transit users is computed by

$$tt_{pt} = \frac{1}{N_{Pt}} \sum_{i=1}^{N_{Pt}} t_i^a - t_i^d, \tag{7.3}$$

where where t_i^a represents the arrival time to its destination and t_i^d is the departure time from the origin of the trip for an agent i. This time includes walking time to the bus stop, waiting time, and bus riding time.

7.2.2 Density

One measure of the effectiveness to define levels of service is density. Density describes the proximity of vehicles to each other, which is the principal influence on freedom to manoeuvre. Thus, it is an appropriate descriptor of service quality(Roess, Prassas, and McShane, 2011). Density is a relationship given by $D = v \ / \ S$ where v is the volume or flow rate, S is the average passenger-car speed(km/h) and D is the density (pc/km/ln)(Roess, Prassas, and McShane, 2011). For our scenario, density is computed by

$$density = \frac{N_{TripsPc} + N_{TripsPt}}{AvgSpeed} \times \frac{1}{AvgNLinks}, \tag{7.4}$$

where $N_{TripsPc}$ and $N_{TripsPt}$ are the total number of trips of private vehicles and public transportation, respectively, computed on the routes taken by agents. *AvgSpeed* is the total average speed of vehicles during the scenario duration time. *AvgNLinks* is the average of links traversed for each agent during its trip.

7.2.3 Fuel Consumption & Emissions

Fuel consumption (fc) of the agents along their legs can be expressed by

$$fc = \sum_{i=1}^{V} \sum_{j=1}^{L} c_i^j \tag{7.5}$$

where V is the number of vehicles (public and private), L the number of legs, and c_i^j is the fuel consumption (in grams/km) of the i^{th} vehicle at the j^{th} leg. Emissions are computed from the output generated by MATSim extension, which is called along with the travel details of the agents produced by MATSim and the profiles of the vehicles.

7.3 Simulation Results and Discussion

7.3.1 Hypervolume

First, we compute the hypervolume of the non-dominated set of solutions at each generation to verify their evolution. The reference point to calculate

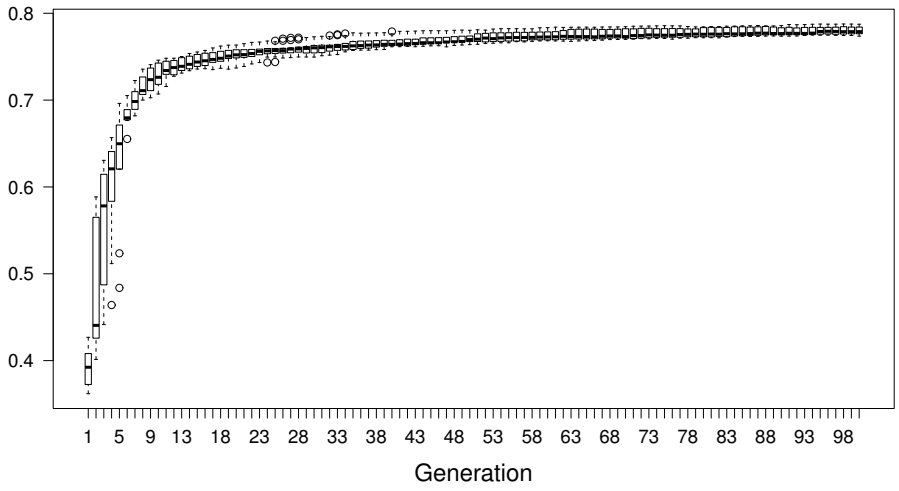


FIGURE 7.1: Hyper Volume Over Generations

the hypervolume is fixed to the worst values in each objective, obtained from the Pareto set of non-dominated solutions found by the algorithm in all runs. Fig. 7.1 shows box plots of the normalized hypervolume over the generations for the ten runs of the algorithm. Note that the hypervolume rapidly increases with the generations and variance reduces considerably by the end of the search.

TABLE 7.1: Objective Correlation Matrix

	tt	dens	fc
tt	—	-0.94	-0.88
dens	-0.94	—	0.84
fc	-0.88	0.84	—

TABLE 7.2: Levels of service for basic freeways segments

Levels of Service	**Density (pc/km/ln)**
A	0-7
B	7-11
C	11-16
D	16-22
E	22-28
F	>28

7.3.2 Trade-offs

Next, we analyze trade-offs between objectives. Unless stated otherwise, we use in our analysis the set of non-dominated solutions found in the ten runs of the algorithm. The correlation values between objectives are shown in Table 7.1, where negative correlations suggest trade-offs between objectives.

Fig. 7.2 shows the projection on the plane *fuel consumption - travel time* of the three-dimensional Pareto front. Solutions in the plot are colored by the level of service (LoS), computed from the third objective density. To compute LoS we use the six levels of service defined by The Highway Capacity Manual (Board, 2000), designated by the letters A through F, with A being the highest level of service and F the lowest. Table 7.2 shows the different levels according to density. In the figure, the colors correspond to LoS A-D.

From Fig. 7.2 note that there is a perceptible trade-off between fuel consumption and travel time, in agreement with the negative correlation shown in Table 7.1. Looking at the levels of service, note that in general solutions with the best LoS, such as A and B, have low values of fuel consumption, but high travel times. This is because high levels of service imply low traffic densities. High levels of service are achieved under scenarios with more public transportation users, where, on average takes the user longer to get to their destinations, but less private vehicles circulate the city. Solutions with worst levels of service, C and D, in general show best travel time but the worst fuel consumption. This is because more people uses their cars instead of public transportation. However, note that within each level of service there also are trade-offs between travel time and fuel consumption. These solutions represent a diversity of configurations, especially in the proportion of public transportation and private car users. To analyze this with more detail, Fig. 7.3 shows the Pareto front colored according to the proportion of public transportation users (N_{Pt}).
Comparing with Fig. 7.2, note that a level of service A requires that 70% and 80% of the agents use public transportation. Level of service B can be achieved with 50% and 60% of these agents. However, level C can span the range from 30% to 50%, and level D requires 20%. This is crucial because it shows that, for the same number of agents that use public transportation, some public transport configurations allow to reduce travel time but could deteriorate the level of service and vice versa. These trade-offs should be studied further to provide useful information to those in charge of planning the transportation services in the city. In Fig. 7.2 and 7.3 also note that the worst level of service is for $N_{Pt} = \mathbf{20\%}$ when more people uses private cars, which is consistent with high densities of traffic as explained below and seen in Fig. 7.4.

Fig. 7.4 shows the trade-off between *travel time* and *density*, which is in agreement with the negative correlation shown in Table 7.1. This figure is colored according to the proportion of agents N_{Pt} that use public transportation. Also, it includes vertical lines to mark the ranges of the levels of service A-D to offer another perspective of the data. From this figure is quite apparent that high densities of traffic correspond to the smaller proportion of agents using public transportation and vice versa.

7.3.3 Emissions

Another important aspect of sustainable public transportation is its environmental impact. As stated above, larger fractions of public transportation

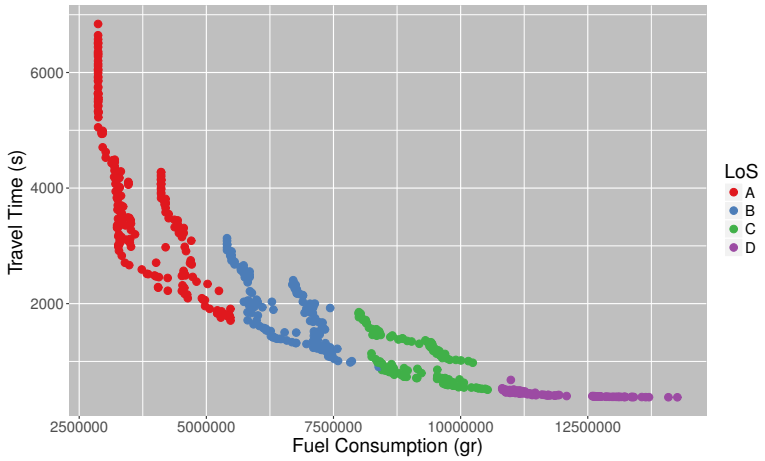


FIGURE 7.2: Fuel consumption and travel time, colored by level of service LoS

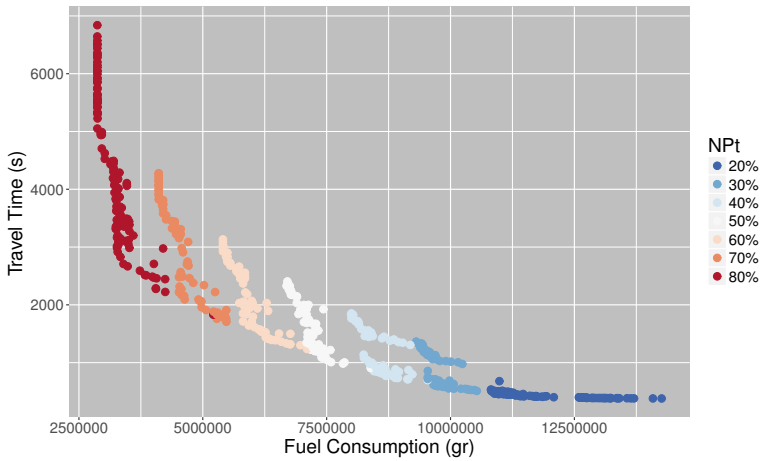


FIGURE 7.3: Fuel consumption and travel time, colored by fraction of public transport users N_{Pt}

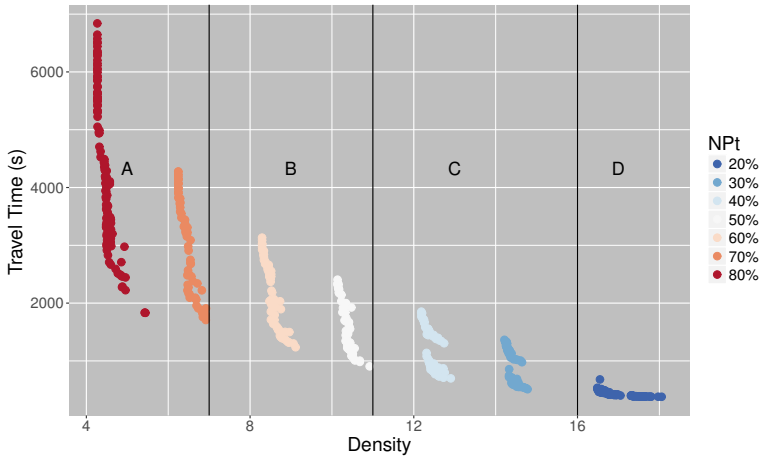


FIGURE 7.4: Travel time and density, colored by fraction of public transport users N_{Pt}

users N_{Pt} reduce overall fuel consumption in the city as shown in Fig. 7.3. Also, best levels of service and low levels of fuel consumption can be achieved simultaneously, as shown in Fig. 7.2.

In addition to fuel consumption, we collect data from MATSim's emissions module. In this work, we analyze particulate matter (PM) emissions. The concentration of PM is composed of engine exhaust gas emissions as well as abrasion and suspension. Regarding health effects, the smaller the particle, the worse the impact on humans. The effects range from damages of the respiratory system to carcinogenic effects.

Fig. 7.5 shows travel time against PM emissions of the non-dominated solutions found by the evolutionary algorithm, coloring by the level of service. In general, it can be seen a trade-off between travel time and PT emissions. Note that PM emissions tend to increase while the level of service decreases from A to D.

Fig. 7.6 and Fig. 7.7 show box plots of the number of public transportation trips and PM emissions grouped by LoS, respectively. From these figures, note that the number of trips and PM emissions increase with density, as the level of service reduces from A to D. In this work the number of buses are not constrained, this explains why there are several solutions with both worst level of service and a large number of bus trips.

To clarify the trade-off between fuel consumption (energy) and PM emissions, we analyze with more detail two solutions relatively close in fuel consumption and travel time, classified with levels of service A and B. One is set with the proportion of public transportation users to $N_{Pt} = 70\%$ and level A and the other one with $N_{Pt} = 60\%$ and level B. We call them A7 and B6, respectively, as shown in Fig. 7.8. Solution B6 achieves slightly lower travel time than A7, but fuel consumption is higher in B6 because there are more private car users. However, note that A7 with lower fuel consumption than B6 produces more PM emissions, as shown in Fig. 7.9. This is because solution A7 has a different headway compared with B6. In A7 midday frequencies for bus lines 2, 3 and 5 are higher than B6.

The geo-located differences in PM emissions between A7 and B6 are illustrated in Fig. 7.10, where red color indicates an increase of PM emissions in A7 and blue color reduction of PM emissions. Note that the levels of PM increase considerably on the network segments where bus lines are located (see Fig. 6.1 for the location of bus lines), where A7 has a high frequency in midday respect to B6. On the other hand, the reduction in emissions shown by the blue color is because A7 has a smaller number of private cars moving in the zone.

7.4 Conclusions and Future Work

This work studied levels of service in urban transportation on a scenario of a real world traffic network. We coupled a multi-objective evolutionary algorithm with the multi-agent transport simulator MATSim to perform a bi-level search of the proportions of private/public transportation users and optimal configurations of buses capacities and headways associated to each proportion. The multi-objective evolutionary algorithm minimized travel time, fuel consumption and density simultaneously. We defined levels of service in terms of density and analyzed it in conjunction with the trade-offs between

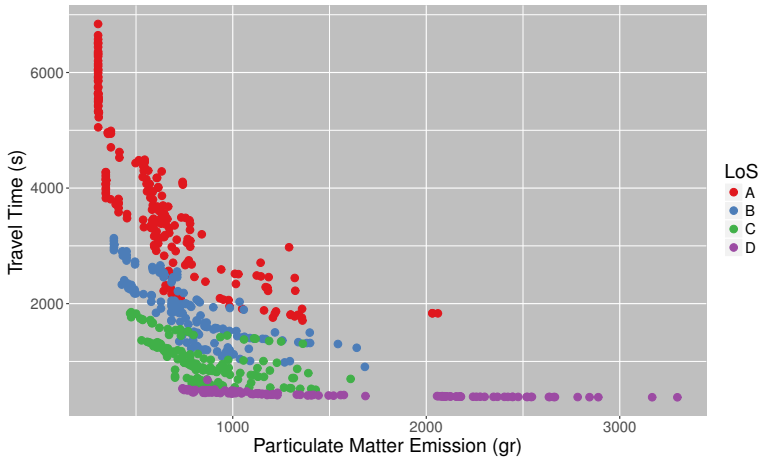


FIGURE 7.5: Trade-off TT and PM (By LoS)

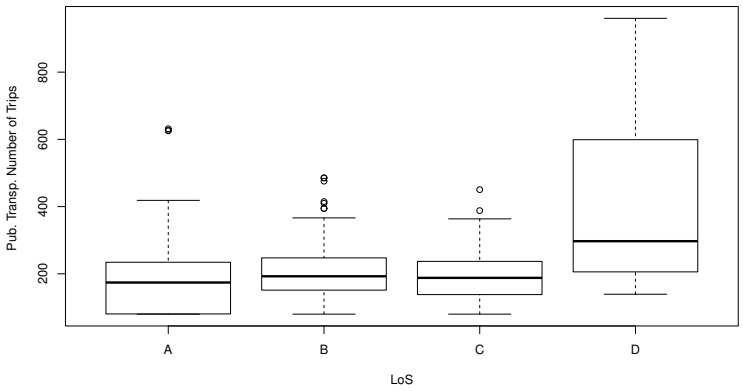


FIGURE 7.6: Number of Public Transportation Trips by LoS

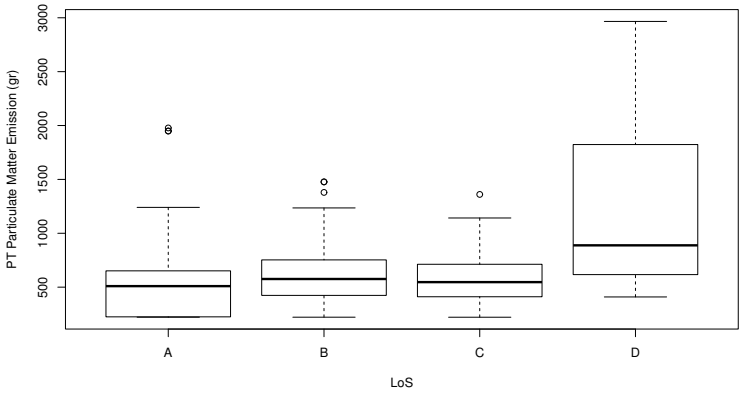


FIGURE 7.7: Public Transportation PM Emission by LoS

objectives on the set of non-dominated solutions found in ten runs of the algorithm. We also analyzed the effects of particulate matter (PM) emission over the area. The results of this study provide valuable insights to understand better the conditions that can lead to different levels of service in the city. It also suggests ways to improve mobility combining private and public

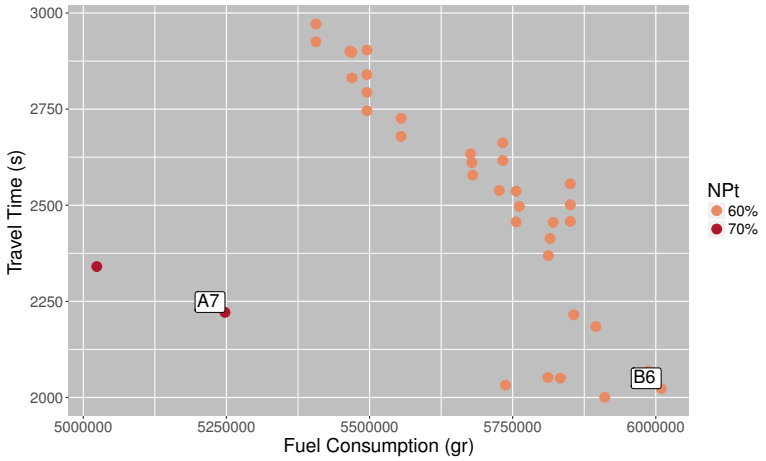


FIGURE 7.8: POS Levels A and B (by N_{Pt})

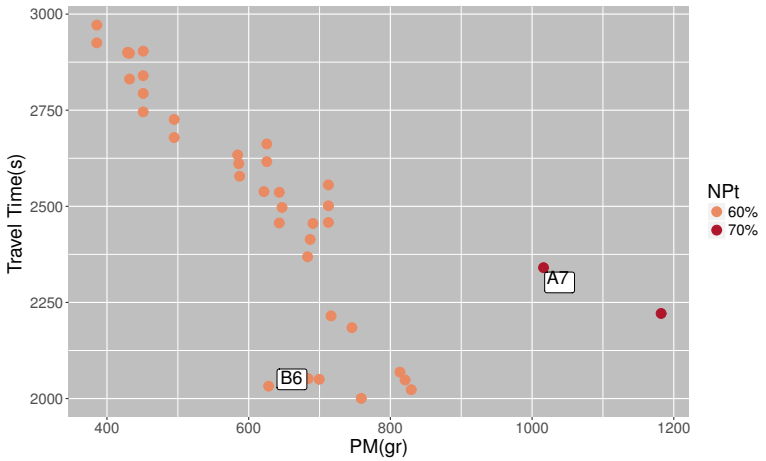


FIGURE 7.9: Travel Time vs PMPT Trade-Off

transportation. Our results, clearly show that to improve LoS, more travelers must use public transportation, but better LoS do not necessary implies a reduction in PM because it depends on the BRT headways.

In the future, we would like to analyze deeper the outcome of the optimization, refine the model to consider constraints in resources, and include other objectives such as cost. We would also like to extend the definition of service to cover other criteria besides density.

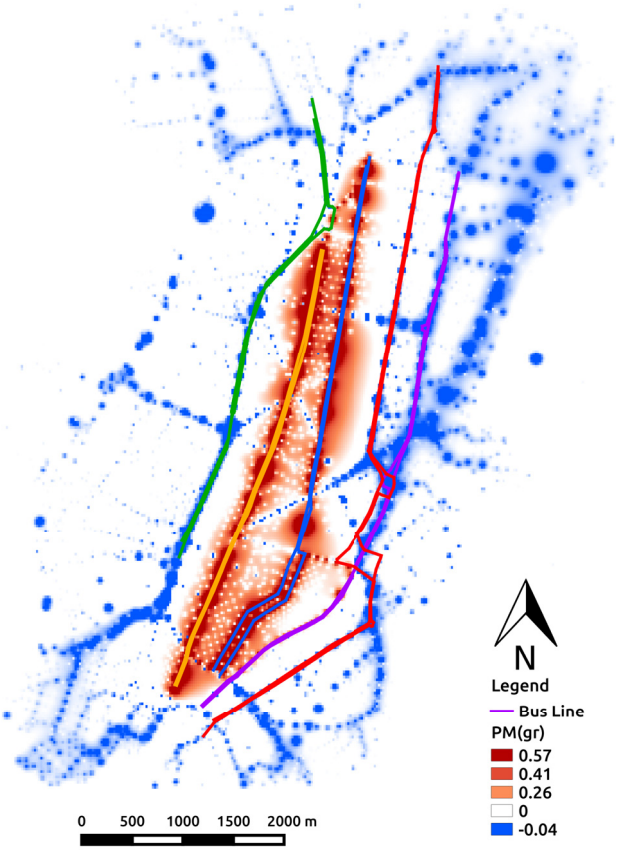


FIGURE 7.10: PM Emission Difference Between A7 and B6

Chapter 8

Conclusions

Some complex problems that involve several variables and their interrelationships and constraints become a continuous challenge since several points of view, but especially since engineering perspective where the expert dealt with some known models and usually he/she does not count with enough resources to explore more alternatives. Mobility and Traffic problems in cities are one of those problems due to not only the complexity but also because the area of study usually is enormous. Additionally, from a sustainable perspective where economic, social and environmental aspects meet, the searching of solutions involves an arduous task to meet simultaneously some objectives that conflict each other. This research has proposed a framework that couples mobility, traffic and emission simulators with single and multi-objective evolutionary algorithms to search and explore alternative designs to mobility and transportation problems. The framework also proposed a method to evolve designs starting from a problem formulation and modelled from a real system, their variables and their respective constraints. Two real problems were considered to test it. One, related to optimization of traffic signals coordination and a second one related to Level of Service optimization. For each problem, a representation and a set of operators were implemented. Quito city was the case of study, where approximately 2 million of people are living and suffer traffic and mobility problems. The area of study focused on the business district area that covers an extension of 40 km^2.

A first problem related to optimization of traffic signal coordination was studied. Seventy signals allocated in the major arteries of the scenario implied not only a huge geographical area but also a vast variable search space. Only a few works previously have dealt with a significant number of signals and variables. To improve the coordination and traffic flow, several genetic operators were implemented. The first set of step operators, one per variable signal (cycle, green time and offset) improved the searching of solutions which was verified by a notable decrease in the vehicles' travel time. The evolutionary algorithm has shown that green times should be mutated with higher probability than cycle and offsets. However, the coordination of the signals was an issue to solve; the step-offset variable operator was modified to consider the propagation of the cycle length to neighbouring signals with offsets set based on distance. It concedes to the vehicles to cross several intersections during the green time phase. Due to the simulated mobility scenario considers a large area of study, each evaluation was computationally

expensive. A deterministic varying mutation was implemented to accelerate the convergence. Also, one and two-point crossover effect was analyzed, with no considerable difference between them, but a better convergence is achieved when it is present in the evolution. The combination of these strategies efficiently and effectively explore the large search space of traffic signal parameters. Moreover, a suitable balance between exploration and exploitation was performed by a combination of a high selection pressure given by elitism with crossover followed by varying mutation. To verify the parameter's effects in the algorithm, an analysis was performed by sequential model-based algorithm configuration (SMAC). The comparison with SMAC configurations showed that varying mutations relatively higher than the recommended in our configuration but using smaller rates of crossover lead to similarly good results. A complete analysis in the optimized solution's variable space confirmed that some patterns emerge. A hierarchical clustering showed that by grouping signals, large chunks of the clusters are spatially contiguous. The geolocalization identified micro-zones with similar characteristics in the city, which is an essential picture for the decision makers to incorporate additional mobility criteria such as safety, emissions and fuel consumption and multi-modality. An examination of emissions revealed that a better optimization process for travel time given by the neighbourhood operators and deterministic varying mutation is also highly correlated to a better optimization for fuel consumptions and emissions. Moreover, a geo-location of emission showed that emissions mostly reduce across the area of study when signals settings are optimized. It provided helpful information for city planners and can be used to feedback the evolutionary algorithm to favour low emissions in certain regions of interest.

A second problem related to optimization of the level of service (LoS) of urban transportation based on traffic density was studied. The study was done in Quito city scenario, modelling the mobility of 27.000 agents that use private and public transportation. Quito's urban transportation system is highly congested due to an increase of private vehicles combined with a poorly satisfied demand by public transportation. The scenario considered five main Bus Rapid Transit (BRT) corridors which are the most demanded and congested routes. A multi-objective evolutionary algorithm was employed to search combinations of a number of private/public transportation users, capacity of buses, and time interval between bus departures. Different proportion of users and configurations will impact travel time, fuel consumption, traffic density, costs, and emissions. Most of these criteria conflict with each other. A first approach optimized density, travel time and fuel consumption simultaneously. The representation considered as variables the proportion of users in public transportation, the capacity and headways for the BRT lines to four schedule segments in the day. Uniform crossover was implemented. The mutation operators implemented, increase or decrease randomly with equal probability the code of the variable in which operates. An analysis of the trade-offs exhibited that in general, solutions with the best level of service, had low values of fuel consumption, but high travel time.

That is because, public transportation users, on average, take the longer to get to their destinations, but less private vehicles circulate the city. Studying the trade-off by the proportion of public transportation users, pointed that, for the same number of agents that use public transportation, some public transport configurations allow to reduce travel time but could deteriorate the level of service and vice versa. Another essential aspect of sustainable public transportation is its environmental impact. A trade-off between particulate matter (PM) emissions and travel time classified by level of service, revealed that PM emissions tend to increase while the level of service decreases. For the decision makers, that kind of analyses could provide useful information to those in charge of planning the transportation services in the city.

A more refined model and bi-objective optimization study were done to examine better the trade-off between travel time and traffic density. The bi-objective formulation allowed to the evolutionary algorithm found some clusters of solutions. A trade-off analysis by the proportion of public transportation users, reported that within each cluster there is no much difference in density, but travel time can change significantly, especially for larger fractions of public transportation users. Clusters also revealed that ranges of improvement in travel time that can be achieved by optimizing public transportation with almost no impact on traffic density for a given value of public transportation users. However, distinguished changes in density were observed between clusters, for different values of public transportation users. That suggests that to reduce density a significant number of users should be encouraged to use public transportation instead of private cars. Unlike the previous study, in this one, fuel consumption and emissions were considered as part of the solution analysis but not as optimized objective. However, a similar trade-off between travel time and fuel consumption, as previous work was found. A geo-location analysis of two solutions revealed that PM emissions from two different configurations in bus capacity and frequency cause different environmental impacts. In general, the results distinctly explained that to improve LoS, more travellers must use public transportation, but an enhancement in LoS does not necessarily imply a reduction in PM because it depends on the BRT headways and technology, especially fuel type.

As future work, it would be worth implement new scenarios considering multi-modal mobility like car, bus, bikes and pedestrians. Propose a new formulation problem to find the balance in the network transportation to support and guarantee the multimodal level of service. Another valuable study could be to consider the technological transition from vehicles based on fuel to electric, and the impact from a sustainable standpoint.

Regarding the simulation model and its expensive computational time, it imposes to implement new methods to deal with that issue. A new step is the development of new heuristics and operators, or the development of surrogate models to improve the simulation performance and evaluation.

Bibliography

Aguirre, H., A. Oyama, and Tanaka K. (2013). "Adaptive ε-Sampling and ε-Hood for Evolutionary Many-Objective Optimization". In: *Evolutionary Multi-Criterion Optimization*. Vol. 7811. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 322–336.

Aguirre, H and K Tanaka (2003). "A Study on the Behavior of Genetic Algorithms on NK-Landscapes: Effects of Selection, Drift, Mutation, and Recombination". In: *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E86-A.9, pp. 2270–2279.

Aguirre, H. and K. Tanaka (2004). "A Study on Parallel Varying Mutation in Deterministic and Self-Adaptative GAs with 0/1 Multiple Knapsack Problems." In: *IPSJ Transactions on Mathematical Modeling and Its Applications* 45, pp. 77–91.

Aguirre, H. et al. (2014a). "An Analysis on Selection for High-Resolution Approximations in Many-Objective Optimization". In: *Parallel Problem Solving from Nature – PPSN XIII*. Vol. 8672. Lecture Notes in Computer Science. Springer International Publishing, pp. 487–497.

Aguirre, H. et al. (2014b). "Extending AεSεH from Many-objective to Multi-objective Optimization". In: *Simulated Evolution and Learning*. Vol. 8886. Lecture Notes in Computer Science. Springer International Publishing, pp. 239–250. DOI: 10.1007/978-3-319-13563-2_21. URL: http://dx.doi.org/10.1007/978-3-319-13563-2_21.

Ambiente, DMQ Secretaria de (2016). *Informe de Calidad del Aire 2016*. URL: www.quitoambiento.gob.ec.

Arbex, Renato Oliveira and Claudio Barbieri da Cunha (2015). "Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm". In: *Transportation Research Part B: Methodological* 81, pp. 355–376. ISSN: 01912615. DOI: 10.1016/j.trb.2015.06.014. URL: http://dx.doi.org/10.1016/j.trb.2015.06.014.

Armas, R., H. Aguirre, and K. Tanaka (2014). "Effects of Mutation and Crossover Operators in the Optimization of Traffic Signal Parameters". In: *The Tenth International Conference on Simulated Evolution and Learning (SEAL)*. Vol. 8886. Lecture Notes in Computer Science. Springer, pp. 167–179.

Armas, R. et al. (2015). "Traffic Signal Optimization: Minimizing Travel Time and Fuel Consumption." In: *Proc. 12th International Conference, Evolution Artificielle (EA'2015), Lyon, France*. Vol. 9554. Lecture Notes in Computer Science. Springer.

Armas, R. et al. (2016a). "An effective EA for short term evolution with small population for traffic signal optimization". In: *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–8. DOI: 10.1109/SSCI.2016.7850099.

Armas, R. et al. (2016b). "Traffic signal optimization and coordination using neighborhood mutation". In: *2016 IEEE Congress on Evolutionary Computation (CEC)*, pp. 395–402. DOI: 10.1109/CEC.2016.7743821.

Back, T. and M. Schutz (1996). "Intelligent Mutation Rate Control in Canonical Genetic Algorithms". In: *Foundations of Intelligent Systems*. Vol. 1079. Lecture Notes on Artificial Intelligence. Springer, pp. 158–167.

Bentley, Peter J. (1999). *Evolutionary Design by Computers*. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. ISBN: 155860605X.

Bielli, Maurizio, Massimiliano Caramia, and Pasquale Carotenuto (2002). "Genetic algorithms in bus network optimization". In: *Transportation Research Part C: Emerging Technologies* 10.1, pp. 19–34. ISSN: 0968090X. DOI: 10.1016/S0968-090X(00)00048-6.

Board, Transportation Research (2000). *Highway Capacity Manual*. Ed. by Washington DC National Research Council. National Research Council,Washington DC.

Castiglione, J. (2015). *Activity-Based Travel Demand Models: A Primer*. Ed. by Transportation Research Board. Transportation Research Board.

Ceylan, H. and M. Bell (2004). "Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing". In: *Transportation Research Part B: Methodological* 38.4, pp. 329–342. ISSN: 01912615. DOI: 10.1016/S0191-2615(03)00015-8.

Chen, J. and L. Xu (2006). "Road-Junction Traffic Signal Timing Optimization by an adaptive Particle Swarm Algorithm". In: *Proc. 9th Int. Conf. Control, Autom., Robot. Vis.* Pp. 1–7.

Coello, Carlos A. Coello, Gary B. Lamont, and David A. Van Veldhuizen (2006). *Evolutionary Algorithms for Solving Multi-Objective Problems (Genetic and Evolutionary Computation)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc. ISBN: 0387332545.

Demoraes, F. (2005). "Movilidad, elementos escenciales y riesgos en el Distrito Metropolitano de Quito". PhD thesis. Universidad de Savoie - Francia.

DMQ (2012). *Encuesta domiciliaria de movilidad (EDM11) del Distrito Metropolitano de Quito (DMQ)*.

— (2015). *Empresa Publica Metropolitana de Transporte de Pasajeros - Distrito Metropolitano de Quito Rendicion de Cuentas 2015*. Document.

Eckart, Z. (2004). "A Tutorial on Evolutionary Multiobjective Optimization". In: *Metaheuristics for Multiobjective Optimisation*.

Ellis, N. (1972). *Driver Expectancy: Definition for Design. Research Report Numb. 606-5 1972*. Tech. rep. https://library.ctr.utexas.edu/digitized/TexasArchive/MS1034.pdf,Accessed: October-2014. Texas Transportation Institute, Texas A&M University.

Farhan Ahmad (2005). "A Genetic Algorithm Based Bus Scheduling Model for Transit Network". In: *Eastern Asia Society for Transportation Studies*. 5, pp. 477–489.

Federal Highway Administration, Department of Transportation (2010). *Traffic Signal Manual*. https://ops.fhwa.dot.gov/publications/fhwahop08024/chapter5.htm, Accessed: October-2014. Federal Highway Administration,

Department of Transportation. URL: https://ops.fhwa.dot.gov/publications/fhwahop08024/chapter5.htm.

Frederik, R., J. Topf, and C. Karch (2007). *Geofabrik*. http://www.geofabrik.de, Accessed: January 2014. URL: http://www.geofabrik.de.

Gakenheimer, R. (1999). "Urban Mobility in the developing world." In: *Transportation Research Part A: Policy and Practice* 33, pp. 671–689.

Garber Nicholas, Hoel Lester (2009). *Traffic & Highway Engineering*. CENGAGE Learning.

Garcia-Nieto, J., E. Alba, and C. Olivera (2012). "Swarm intelligence for traffic light scheduling: Application to real urban areas". In: *Engineering Applications of Artificial Intelligence* 25.2, pp. 274–283.

Garcia-Nieto, J., C. Olivera, and E. Alba (2013). "Optimal cycle program of traffic lights with particle swarm optimization". In: *IEEE Transactions on Evolutionary Computation* 17.6, pp. 823–839. ISSN: 1089778X. DOI: 10.1109/TEVC.2013.2260755.

Grether, D. and A. Neumann (2011). *Traffic Light Control in Multi-Agent Transport Simulations*. Tech. rep. Transport Systems Planning and Transport Telematics, Technical University Berlin.

GrupoFaro (2010). *Reverdeciendo las politicas: Hacia una movilidad sustentable en Quito: El potencial de la bicicleta como medio de transporte alternativo*. Espanol. Accessed: January 2017. Grupo Faro. URL: http://www.grupofaro.org/.

G.Scora and M. Barth (2006). *Comprehensive Modal Emission Model (CMEM), version 3.01 User's Guide*. University of California Riverside Center for Environmental Research and Technology.

Hee, Kim Jin, Bae Yun Kyung, and Chung Jin Hyuk (2012). "Multi objective Optimization for Sustainable Road Network Design Problem". In: *International Conference on Transport, Environment and Civil Engineering (ICTECE'2012)*, p. 5.

Hong, Y.S. et al. (1999). "Estimation of Optimal Green Time Simulation Using Fuzzy Neural Network". In: *Fuzzy Systems Conference*, pp. 761–766.

Horni, A., K. Nagel, and K.W. Axhausen (2016). *The Multi-Agent Transport Simulation MATSim*. English. http://matsim.org, Accessed: January 2014. Ubiquity, London. URL: matsim.org.

Hülsmann, F et al. (2011). "Towards a multi-agent based modeling approach for air pollutants in urban regions." In: *Proceedings of the Conference on "Luftqualität an Straßen"* FGSV Verlag GmbH, pp. 144–166. URL: http://www.matsim.org/extension/emissions.

Hutter, F., H. H. Hoos, and K. Leyton-Brown (2011). "Sequential Model-Based Optimization for General Algorithm Configuration". In: *Proceedings of the 5th International Conference on Learning and Intelligent Optimization (LION-5)*. Springer, pp. 507–523.

INEC (2010a). *El Transporte Terrestre de Pasajeros en Ecuador y Quito: Perspectiva Histórica y Situación Actual*.

— (2010b). *Equadorian National Institute of Statistics*. http://www.ecuadorencifras.gob.ec/, Accessed: October-2014.

Kachroudi, S. and N. Bhouri (2009). "A multimodal traffic responsive strategy using particle swarm optimization". In: *In Proc. 12th IFAC Symposium on Control in Transportation Systems*. Vol. 42. 15, pp. 531–537.

Keller, M. and P. Wuthrich (2014). *Handbook emission factors for road transport 3.1 / 3.2*.

Kickhofer, B. (2014). "Economic Policy Appraisal and Heterogeneous Users". PhD thesis. Technische Universitat Berlin. DOI: http://dx.doi.org/10.14279/depositonce-4089.

Kickhöfer, B. et al. (2013). "Rising car user costs: comparing aggregated and geo-spatial impacts on travel demand and air pollutant emissions." In: *Smart Transport Networks: Decision Making, Sustainability and Market structure*, pp. 180–207.

Knowles, Joshua D., Richard A. Watson, and David Corne (2001). "Reducing Local Optima in Single-Objective Problems by Multi-objectivization". In: *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization*. EMO '01. London, UK, UK: Springer-Verlag, pp. 269–283. ISBN: 3-540-41745-1. URL: http://dl.acm.org/citation.cfm?id=647889.736521.

Laumanns, Marco et al. (2002). "Combining Convergence and Diversity in Evolutionary Multiobjective Optimization". In: *Evolutionary Computation* 10.3, pp. 263–282. DOI: 10.1162/106365602760234108. URL: http://dx.doi.org/10.1162/106365602760234108.

Ma, Jihui et al. (2017). "A Model for the Stop Planning and Timetables of Customized Buses". In: pp. 1–28. DOI: 10.1371/journal.pone.0168762.

McKenney, D. and T. White (2013). "Distributed and adaptive traffic signal control within a realistic traffic simulation". In: *Engineering Applications of Artificial Intelligence* 26.1, pp. 574–583.

Michalewicz, Z. and D.B. Fogel (2010). *How to Solve It: Modern Heuristics*. Springer Berlin Heidelberg. ISBN: 9783642061349. URL: https://books.google.co.jp/books?id=Zt33kQAACAAJ.

Miller, A. J. (1963). "Settings for fixed-cycle traffic signals". In: *Operational Research Quarterly* vol. 14.no. 4, pp. 373–386.

Murtagh, F. and P. Legendre (2014). "Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?" In: *Journal of Classification* 31.3, pp. 274–295. ISSN: 1432-1343. DOI: 10.1007/s00357-014-9161-z. URL: http://dx.doi.org/10.1007/s00357-014-9161-z.

OpenStreetMap. *Osmosis*. English. http://wiki.openstreetmap.org/wiki/Osmosis Accessed: May 2013. URL: https://github.com/openstreetmap/osmosis.

Park, B., C. Messer, and T. Urbanik (1999). "Traffic Signal Optimization Program for Oversaturated Conditions: Genetic Algorithm Approach". In: *Transportation Research Record* 1683.1, pp. 133–142. ISSN: 0361-1981. DOI: 10.3141/1683-17.

Peng, L. et al. (2009). "Isolation niches particle swarm optimization applied to traffic lights controlling". Undetermined. In: *In Proc. of the 48h IEEE Conference on Decision and Control (CDC)*. IEEE, pp. 3318–3322. DOI: 10.1109/CDC.2009.5399767.

QGIS Development Team (2009). *QGIS Geographic Information System*. Open Source Geospatial Foundation. URL: http://qgis.osgeo.org.

Reed, P.M. et al. (2013). "Evolutionary multiobjective optimization in water resources: The past, present, and future". In: *Advances in Water Resources* 51.Supplement C. 35th Year Anniversary Issue, pp. 438 –456. ISSN: 0309-1708. DOI: https://doi.org/10.1016/j.advwatres.2012.01.005. URL: http://www.sciencedirect.com/science/article/pii/S0309170812000073.

Richardson, B. C. (2005). "Sustainable transport: analysis frameworks". In: *Journal of Transport Geography* 13.1, pp. 29–39. ISSN: 0966-6923.

Roess, R. P., E. Prassas, and W. McShane (2011). *Traffic Engineering*. Ed. by Pearson Prentice Hall. Fourth. Pearson Education.

Rouphail, N., B. Park, and J. Sacks (2000). "Direct Signal Timing Optimization: Strategy Development and Results". In: *Proc. XI Panamerican Conference in Traffic and Transport Engineering*. 9, pp. 195–206.

Sánchez, J., M. Galán, and E. Rubio (2008). "Applying a Traffic Lights Evolutionary OptimizationTechnique to a Real Case: " Las Ramblas " Area inSanta Cruz de Tenerife". In: *IEEE Transactions on Evolutionary Computation* 12.1, pp. 25–40.

— (2010). "Traffic Signal Optimization in "La Almozara" District in Saragossa Under Congestion Conditions, Using Genetic Algorithms, Traffic Microsimulation, and Cluster Computing". In: *IEEE Transactions on Intelligent Transportation Systems* 11.1, pp. 1–10.

Shepard, D. (1968). "A Two-dimensional Interpolation Function for Irregularly-spaced Data". In: *Proceedings of the 1968 23rd ACM National Conference*. ACM '68. New York, NY, USA: ACM, pp. 517–524. DOI: 10.1145/800186.810616. URL: http://doi.acm.org/10.1145/800186.810616.

Spears, W. (2000). *Evolutionary Algorithms: The Role of Mutation and Recombination*. Springer-Verlag Berlin.

Spiegelman, C., E. Sug-Park, and L. Rilett (2011). *Transportation Statistics and Microsimulation*. CRC Press.

Srinivasan, D., M. C. Choy, and R. Cheu (2006). "Neural netwokrs for real time traffic signal control". In: *IEEE Transactions on Intelligent Transportation Systems* 7.3, pp. 261–272.

Stevanovic, Aleksandar et al. (2011). "Traffic control optimization for multimodal operations in a large-scale urban network". In: *2011 IEEE Forum on Integrated and Sustainable Transportation Systems, FISTS 2011*, pp. 146–151. ISBN: 9781457709906. DOI: 10.1109/FISTS.2011.5973603.

STM (2008). *Signal Timming Manual*. Federal Highway Administration, USA.

Sugiyama, Yuki et al. (2008). "Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam". In: *New Journal of Physics* 10.3, p. 033001. URL: http://stacks.iop.org/1367-2630/10/i=3/a=033001.

Suzuki, R. and H. Shimodaira (2015). *pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling*. URL: https://CRAN.R-project.org/package=pvclust.

Taniguchi, E. and H. Shimamoto (2004). "Intelligent transportation system based dynamic vehicle routing and scheduling with variable travel times". In: *Transportation Research Part C: Emerging Technologies* 12.3, pp. 235–250.

Teklu, F. (2006). "A Genetic Algorithm Approach for Optimizing Traffic Control Signals Considering Routing". In: *Computer Aided Civil and Infrastructure Engineering* 22, pp. 31–43.

Tomforde, S et al. (2008). "Decentralised Progressive Signal Systems for Organic Traffic Control". In: *IEEE International Conference on Self-Adaptive and Self-Organizing Systems*, pp. 413–422.

Turky, A. M. et al. (2009). "Using Genetic Algorithm for Traffic Light Control System with a Pedestrian Crossing". In: *Proc. Rough Sets and Knowledge Technology (RSKT): 4th International Conference*. Springer, pp. 512–519.

United Nations, Department of Economic and Population Division Social Affairs (2014). *World Urbanization Prospects: The 2014 Revision*.

Univ, Shinshu Univ-Lille. *MODO Frontiers in Massive Optimization and Computational Intelligence*. English. Accessed: January 2016. URL: https://sites.google.com/view/lia-modo/home?authuser=0.

Verbas, İ. Ömer and Hani S. Mahmassani (2015). "Integrated Frequency Allocation and User Assignment in Multimodal Transit Networks". In: *Transportation Research Record: Journal of the Transportation Research Board* 2498, pp. 37–45. ISSN: 0361-1981. DOI: 10.3141/2498-05.

Ward, J. H. (1963). "Hierarchical Grouping to Optimize an Objective Function". In: *Journal of The American Statistical Association* 58 (301), pp. 236–244. DOI: 10.1080/01621459.1963.10500845.

Webster, F. V. (1958). "Traffic signal setting". In: *Road Res. Lab., HMSO, London, U.K., Tech.* Paper 39, pp. 1–44.

Yu, Bin, Zhongzhen Yang, and Jinbao Yao (2010). "Genetic Algorithm for Bus Frequency Optimization". In: *Journal of Transportation Engineering* 136.6, pp. 576–583. ISSN: 0733-947X. DOI: 10.1061/(ASCE)TE.1943-5436.0000119. URL: http://ascelibrary.org/doi/10.1061/{\%}28ASCE{\%}29TE.1943-5436.0000119.

Zhao, D., Y. Dai, and Z. Zhang (2012). "Computational Intelligence in Urban Traffic Signal Control: A Survey". In: *IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews)* 42.4, pp. 485–494.

Publications

Journals

1. R. Armas, H. Aguirre, F. Daolio, K. Tanaka, *Evolutionary Design Optimization of Traffic Signals Applied to Quito City*, Public Library of Science – PLOSONE, https://doi.org/10.1371/journal.pone.0188757, December 2017.

International Conferences

1. R. Armas, H. Aguirre and K. Tanaka *Effects of Mutation and Crossover Operators in the Optimization of Traffic Signal Parameters*. In: Dick G. et al. (eds) Simulated Evolution and Learning. SEAL 2014. Lecture Notes in Computer Science, vol 8886. Springer.

2. R. Armas, H. Aguirre, S. Zapotecas-Martinez and K. Tanaka *Traffic Signal Optimization: Minimizing Travel Time and Fuel Consumption*. In: Bonnevay S., Legrand P., Monmarché N., Lutton E., Schoenauer M. (eds) Artificial Evolution. EA 2015. Lecture Notes in Computer Science, vol 9554. Springer.

3. R. Armas, H. Aguirre, F. Daolio and K. Tanaka, *Traffic signal optimization and coordination using neighborhood mutation*, 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, BC, 2016, pp. 395-402. doi: 10.1109/CEC.2016.7743821

4. D. Stolfi, R. Armas, E. Alba, H. Aguirre, K. Tanaka, *Fine Tuning of Traffic in our Cities with Smart Panels: The Quito City Case Study*, The Genetic and Evolutionary Computation Conference (GECCO), Denver, 2016.

5. R. Armas, H. Aguirre, F. Daolio and K. Tanaka, *An effective EA for short term evolution with small population for traffic signal optimization*, 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, 2016, pp. 1-8. doi: 10.1109/SSCI.2016.7850099

6. R. Armas, H. Aguirre and K. Tanaka, *Multi-Objective Optimization of Level of Service in Urban Transportation*, The Genetic and Evolutionary Computation Conference (GECCO), Berlin, 2017.

Book Chapter

1. R. Armas, and H. Aguirre, 2016. *Quito Metropolitan District*. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) The Multi-Agent Transport Simulation MATSim, pp. 473-476. London: Ubiquity Press. License: CC-BY 4.0

Local Conferences (Japan)

1. R. Armas, H. Aguirre and K. Tanaka, *A Fast and Reliable EA for Traffic Signal Optimization*, 2016 IEEE Shin-etsu Session.

2. R. Armas, H. Aguirre and K. Tanaka, *Public Transportation Multi-Objective Optimization for Various Demand Ratios*, 2017 IEEE Shin-etsu Session.