



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

FISICA

Ciclo 37

Settore Concorsuale: 01/A4 - FISICA MATEMATICA

Settore Scientifico Disciplinare: MAT/07 - FISICA MATEMATICA

EMERGENT PATTERNS IN HUMAN MOBILITY: A STATISTICAL PHYSICS
APPROACH

Presentata da: Alberto Amaduzzi

Coordinatore Dottorato

Alessandro Gabrielli

Supervisore

Armando Bazzani

Co-supervisore

Daniel Remondini

Riccardo Gallotti

Esame finale anno 2026

Emergent patterns in human mobility: A statistical physics approach

January 13, 2026

Abstract

The rapid development of Information and Communication Technologies (ICT) has opened up new possibilities for the realization of future smart cities.

We view smart cities as rich and heterogeneous systems composed of a multitude of microscopic, dynamic, and interacting entities. These range from individuals to vehicles, whose collective behavior gives rise to complex macroscopic patterns.

Smart cities face numerous challenges, including the sustainability of tourism, traffic congestion and associated pollution, and the spread of diseases. Crucially, these issues are deeply intertwined with human mobility at various spatial and temporal scales.

A central theme of this thesis is the study of emergent properties—macroscopic phenomena that arise from the interaction of microscopic agents. One such example is traffic, which emerges from the individual mobility demand, the individual behavior in the use of transport networks and from the vehicle dynamics on the road network.

To address these challenges, the physics of complex systems aims to highlight universal properties of emergent phenomena using the data analysis and the statistical mechanics methods. On the one hand data analysis provides the empirical basis for modeling, on the other statistical physics offers a natural framework to abstract away microscopic details and define collective macroscopic observables that capture system-wide behavior.

The overarching goal of this thesis is to characterize emergent properties of mobility systems using microscopic dynamical data on individual behavior.

In the first part of our work, we study GPS data aggregated at the country level—without detailed traffic resolution—to characterize differences in mobility behavior between residents of California and Massachusetts. We assess the sensitivity of our findings across different temporal resolutions. In the second part we analyze GPS data at road-level resolution to study multimodal characteristics of urban traffic. We develop an algorithm to classify trajectories into distinct traffic regimes and to identify sub-networks affected by congestion. This analysis reveals empirical statistical laws, suggesting universal features of human mobility in Bologna and highlight a new measure to quantify traffic in the identified sub-networks. We further test the robustness of these laws across different periods, including holidays.

Using a GPU we simulate traffic to investigate car dynamics under varying morphological configurations (e.g., monocentric vs polycentric origin-destination patterns). Our goal is to identify critical transitions in traffic flow, particularly the onset of congestion, by analyzing the fraction of agents who become "trapped" in the network.

Finally, we analyze an agent-based model of pedestrian tourist flows in Venice. The model is designed to generate mobility patterns comparable to data collected from road-mounted cameras, enabling validation and further exploration of pedestrian dynamics in a city heavily impacted by tourism.

Contents

0.1	Guide to the reader	3
1	Introduction	4
1.1	General Motivations	4
1.2	Mobility changes during Covid-19	5
1.3	Traffic and multilayer structure of urban road networks	7
1.4	Traffic phase transition	8
1.5	Venice	10
2	Data	11
2.1	Data for Mobility changes during Covid-19	13
2.1.1	Data	13
2.1.2	Notations	13
2.2	MDT Data on the metropolitan area of Bologna	15
2.2.1	Data	15
2.2.2	Notations	16
2.3	Facebook population data for Traffic phase transition	19
2.3.1	Data	19
2.3.2	Notations	21
2.4	Pedestrian flow dataset in Venice	23
2.4.1	Data	23
3	Methods and Models	24
3.1	Mobility changes during Covid-19	24
3.1.1	Metrics	24
3.2	Traffic multimodality in Bologna	26
3.2.1	Constructing road network	26
3.2.2	Reconstruction of Trajectories with road network granularity	27
3.2.3	Filtering Trajectories	27
3.2.4	georeferencing and trajectory reconstruction	29
3.2.5	Fuzzy Algorithm: clustering paths	31
3.2.6	Reconstruction of multilayer character of road network	31
3.2.7	Hierarchical re-assignment class	37
3.3	Traffic phase transition	38
3.3.1	Traffic model and phase transition	38
3.3.2	Traffic Simulator	40
3.3.3	Monocentrism/Polycentrism	40
3.3.4	Directed percolation	43
3.3.5	Gravitational Model for flows	45
3.3.6	Changing the shape of the city: Generation of new flows	45
3.4	Pedestrian mobility in Venice	48

3.4.1	Simulator pedestrian mobility in Venice	48
3.4.2	Ward Algorithm	50
3.5	Another Method to Measure Monocentrism	50
3.5.1	Theoretical Framework: from pairwise votes to ranking	51
3.5.2	Definitions	52
3.5.3	Main Theorem	53
3.5.4	Vector Field	55
3.5.5	Potential	55
3.5.6	Vector Field Potential and Pairwise Hierarchies	56
4	Results	57
4.1	Mobility changes during COVID-19	57
4.1.1	California	57
4.1.2	Massachusetts	62
4.2	The multimodal structure of the road network in Bologna	65
4.2.1	Statistical properties paths	65
4.2.2	Speed Distribution and Multi-Modality Network	69
4.2.3	Sub-Networks	74
4.3	Traffic phase transition	75
4.4	Pedestrian mobility in Venice	82
4.5	Field theory for Monocentrism	87
5	Conclusions	91
5.1	Conclusion	91
5.1.1	Mobility changes during COVID-19	91
5.1.2	Characterization of the multimodal structure of the road network in Bologna	92
5.1.3	Traffic Phase Transition	94
5.1.4	Pedestrian mobility in Venice	94
5.1.5	Another Method to Measure Monocentrism	95
A	Appendix	96
A.1	Mobility changes during COVID-19	96
A.2	Characterization of the multimodal structure of the road network	99
A.3	Traffic phase transition	103
A.3.1	Fitting procedure	103
A.3.2	Transition Related Measures	104
A.3.3	Population Distributions	153
A.4	Thanks	159

0.1 Guide to the reader

This thesis is organized in a modular structure around four main studies, each comprising an introduction, data, methods, results, and conclusions section. In addition, it includes a theoretical proposal for a novel measure of monocentrism, which was motivated by one of the four studies and therefore consists only of methods, results, and conclusions.

The work begins with a general introduction outlining the common context across all studies (1.1), followed by a data section of general relevance (2).

The first study examines differences in mobility patterns across periods related to the COVID-19 pandemic. Its objectives are introduced in 1.2, with data and notation presented in 2.1.1. The methodological framework, including the measures applied, is detailed in 3.1. Results are reported in 4.1, with separate analyses for California (4.1.2) and Massachusetts (4.1.2). The conclusions, summarizing the findings and reflecting on their implications, are presented in 5.1.1. Appendix A.1 discusses the robustness of the results under different preprocessing parameters.

The second study investigates congestion properties in Bologna using GPS trajectory data. The introduction and objectives are given in 1.3, with the data and notation described in 2.2.1. The methodological section 3.2 presents the preprocessing steps and theoretical approach, while the results and conclusions are reported in 4.2 and 5.1.2, respectively. Appendix A.2 contains disaggregated daily analyses that complement the average results in the main text and illustrate the stability of the methods under heterogeneous mobility conditions.

The third study addresses the relationship between a phase transition in severely congested road networks and the morphology of origin–destination matrices. The objectives are outlined in 1.4, the data and notation in 2.3.1, and the methodological framework in 3.3. Results are presented in 4.3, followed by conclusions in 5.1.3. Additional material is included in the appendices: the fitting procedures in A.3.1 and the realizations of the phase transition across different morphologies in A.3 and A.3.3.

The fourth study concerns the calibration of a pedestrian simulator for tourist mobility in Venice. Its objectives are presented in 1.5, the data in 2.4.1, and the methodological framework in 3.4. Results and conclusions are reported in 4.4 and 5.1.4, respectively.

Finally, the thesis introduces a theoretical contribution: a novel method for quantifying monocentrism through a potential-based measure. The motivation and methodological development are presented in 3.5, results are discussed in 4.5, and conclusions are provided in 5.1.5.

A note on notation: while each study necessarily adopts conventions tailored to its specific context, the terminology has been kept as intuitive and as homogeneous as possible across the thesis. Besides this, we point out that in each methods section the goals and motivations are restated. At the time of writing, none of the studies has yet been published in peer-reviewed journals.

Chapter 1

Introduction

1.1 General Motivations

The pervasive integration of technology into the lives of people has facilitated the extraction of vast quantities of data, offering insights into various aspects of human existence. One of these facet is human mobility that has been chosen by different authors as paradigmatic example of complexity [51, 31]. Concurrently, a global initiative has emerged— the conception of digital twins for cities [116, 18, 123, 97]. This concept is instrumental in comprehending and replicating urban environments in a digital realm. The overarching goal of this project is to build a system of connected models capable of capturing the complexity of urban dynamics across multiple spatial and temporal scales and their interconnections. As an example could be chosen the problem of understanding how local perturbations in some aspect of mobility (closing a road) , will affect not only local traffic but also public mobility (vehicle delays) and pedestrian mobility (e.g. accessibility to services and housing) Modeling efforts have already been undertaken in various domains, addressing a broad range of urban challenges—from traffic congestion and pollution e.g., [52], [86] to less tangible yet equally important issues such as social segregation [93] and urban safety [[37]]. Listing all the different factors at play when describing urban systems at different scales (individual mobility, road dynamics ecc.) cannot be completely known. This kind of systems can hardly be completely understood from a particular disciplinary perspective, and probably the very finding of a single, consistent, complete and correct model for them is unattainable due to considerations of randomness, spatial or network interdependencies, and nonlinear feedback effects that are difficult to spot and correctly reproduce and calibrate [53]. However studying statistical properties of aggregated communities can highlight some patterns and define surprising levels of predictability [113]. For these reason a lot of different models have been put forward to describe statistical properties of human behavior at short, medium, large space-time granularities [12]. The strength on this approach is the effectiveness of the theories that do not depend on the specific details of what is really going on (the aspects we cannot access). The goal for this thesis is therefore to contribute in this interdisciplinary effort by exploring the statistical properties of human mobility, both considering the dynamics at the street level and the behavioral patterns for large populations and test the stability in time to perturbations like Covid-19 (at National scale) or like holidays and subject to event at traffic scale. Our approach is realized on the base of private datasets (that can be shared just once aggregated), so it is subject to the criticalities described by [122] for reproducibility of results due to privacy laws, and this makes harder to scale the project around the world. The thesis uses statistical approaches for the various problems faced. Each approach stands by itself and this is reflected in the structure of chapters and sections. The first problem we present is the characterization of the change in mobility patterns due to the unprecedented COVID-19 pandemic. This is an example of probing and measuring distributions out of equilibrium at

coarse grained and collective scale. The motivations and the context for this study is introduced in [section 1.2](#), describe the data used in [subsection 2.1.1](#), the methods in [section 3.1](#) and the results in [section 4.1](#).

We then propose a statistical study case for traffic system of the Urban Area of Bologna for different days. Using a framework to distinguish different modalities of using the road network, we study the relation among the multilayer structure of the urban road network and the individual behaviour in realizing their mobility. In this way we discover the spatial scale of the mobility demand in Bologna and we study the existence of universal statistical laws for the mobility paths. We propose a definition of a traffic congestion measure related to the data set that allows us to distinguish congestion during the rush hours in the morning and the afternoon in subsets of the network. We introduce the motivation and the context in [section 1.3](#), describe the data used in [subsection 2.2.1](#), the methods in [section 3.2](#) and the results in [section 4.2](#). We further use two different simulators for pedestrian and car mobility. The two simulators are developed to address the problem of pedestrian mobility in the city of Venice and the microscopic dynamics of regional car mobility. On the one hand we use the simulator of microscopic dynamics of regional car mobility, to shine some light on the relationship between the "morphology" of the area and a specific phase transition [\[49\]](#) to congestion in road networks. We study this transition for different regional areas. We introduce the motivation and the context in [section 1.4](#), describe the data used in [subsection 2.3.1](#), the methods in [section 3.3](#) and the results in [section 4.3](#).

The problem that we approach with the pedestrian simulator is the reproducibility of touristic flows in the city of Venice. We compare flows of people measured from flow sensors mounted in the city of Venice, to the simulator, grouping subsets of cameras according to their correlation matrices. We describe the data used in [subsection 2.4.1](#), the methods in [section 3.4](#) and the results in [section 4.4](#).

1.2 Mobility changes during Covid-19

This section refers to the context and the Facing the unprecedented COVID-19 pandemic raised questions on how to contain the advance of this new disease [\[98, 57\]](#) and their collateral effects [\[100\]](#). In the absence of vaccinations, Non-Pharmaceutical Interventions (NPIs) were the only solutions enforced by governments in the first months of the pandemic with the intent to contain the advance of the disease and played a critical role ([\[3, 73, 43\]](#)). Lockdown and mobility restriction measures were implemented all over the world, and as a direct consequence, human behavior and mobility habits were significantly altered [\[25\]](#). In this context, many models have been built to forecast the spread of COVID-19 and diseases among different populations ([\[29, 124, 50, 67, 30\]](#)). Seminal work on human mobility has shown that aggregate and (pseudo-)anonymized mobile phone data can assist the modeling of the geographical spread of epidemics ([\[115, 119\]](#)), and individual mobility data inferred from mobile phone records have been shown to be instrumental in modeling the spatial spread of epidemics [\[38, 98\]](#). Even more precise estimates of contact patterns have been obtained from GPS trajectories and data provided by location service providers [\[4, 29, 28, 70\]](#). However, the available data often represent only a partial sample of the population, or are available solely in aggregated form or at a spatial granularity that differs from the desired level [\[41, 20, 10\]](#). Consequently, in studies of spreading processes, the use of real-world data is frequently complemented by probabilistic assumptions or mobility models designed to estimate mobility flows between locations and the corresponding probability of contacts [\[6, 8, 115, 32\]](#). The validity of these models, however, may be compromised by the altered characteristics of mobility patterns following the implementation of mobility restrictions [\[81, 107\]](#), and conservative approaches that simply rescale mobility levels may prove inadequate. This study aims to explore how metrics quantifying individual mobil-

ity patterns vary in the presence of stringent mobility restrictions, such as those implemented during the COVID-19 pandemic.

Until now, most studies relied on the assumption of a relationship between the number of trips and the number of contacts, and between the number of contacts and the probability of transmission. Researchers found that this behavior did not remain constant among the different phases of the pandemic, shaping differently according to the stage and driven by awareness ([96, 55]). Therefore deriving information about virus spread from what is known about human mobility remains a challenge [4]. Despite these difficulties, this scenario provides not only the opportunity for developing a quantitative description of the effect on mobility. It also represents a natural experiment that allows us to test the robustness of individual mobility patterns in response to restrictions implemented to limit contact between individuals and see what are the possible factors that have the potential to reshape them. In this work, we aim to quantify these effects by carrying out measures regarding individual mobility. We analyze a large privacy-enhanced longitudinal GPS dataset, including over 180,000 anonymous opted-in individual trajectories of people living in Massachusetts, and more than 1,000,000 from California, from January to September 2020, replicating the analysis from previous seminal works on the statistical description of human mobility. We divide the total analysis into three different periods according to stay-at-home orders [2]. From 15 January to 15 March (before orders were implemented), 15 March to 15 May (when more than 40 states had issued some form of shelter-in-place order), from 15 May to the end of September (corresponding with spring-summer when orders were lifted). For starters, a widely used geometrical quantity measuring how far from some central place a person is willing to move is the radius of gyration. We test, in particular, the effects that political color and the urbanity area play in this measure. This is accompanied by the geodesic distance, which usually is distributed according to power law ([48],[23]). Another measure of interest is the exploration curve that indicates the rate of discovery of new places, and it is well fitted by positive power laws describing diffusive behavior [112]. To describe how people occupy the area around themselves in a comparable way, we compute the distribution of places visited, observed in the intrinsic reference frame [48]. We further put our attention on the predictability of the trajectories. We use three different entropy measures: the Random entropy, which depends on the number of visited places; the Shannon, which adds the dependency on the relative frequency of visits, assuming subsequent visits are independent; and in the end Lempel-Ziv estimation (Real entropy), which also considers the order in which visits are done [113]. These different measures allow a progressively finer view of the trajectories, adding different aspects that contribute to the description of the uncertainty of a single path. Finally, we propose returner-explorers analysis [101] to see how the fraction of most visited places describe the overall tendency to travel further from the center of mass. This measure indicates how much exploration is described by the most visited places. The measure is the fraction between the radius of gyration and the k-radius of gyration (the R_g one has taking into account the first k most visited places).

The work consists of an exploratory analysis of the change in mobility between pre-during-post restrictions during in California and Massachussets This exploration can be a starting point to have a better view on the properties of mobility at population level that can be in principle used to produce synthetic trajectories for statistical modelling of transitions between different meta-populations.

1.3 Traffic and multilayer structure of urban road networks

Urban traffic is a paradigmatic example of complex system. To provide possible scenarios researchers have developed models of increasing complexity to understand the relation between individual microscopic dynamics and the emergent state of traffic. On the one hand we have Lagrangian models [95, 69, 71] that describe the dynamics of single particles (vehicles). These models give a reasonable description of the microscopic dynamic but need lots of data and assumptions to run realistic traffic whose validation is still an open problem. On the other hand we have Eulerian description of the dynamics at network. Such models offer a lightweight possibility of description. In particular fluid-dynamic models [64, 68], offer the possibility to study emergent properties of urban mobility like velocity distribution, fluxes and density of cars at road level even if they loose the granular character of traffic. Of particular interest, due to the predicted stationarity, is the so called Macroscopic Fundamental diagram (MFD) that relates density and fluxes and quantifies a congestion degree in the city. The conditions for the existence of a MFD for an urban road network requires to consider homogeneous roads [34] that is not the case in realistic situations. For a long time results could not be tested due to lack of data [54], but lately thanks to the work of [42], some evidence of such fundamental diagram has been found. This measure allows to define congestion as the state for which, when increasing the number of vehicles decreases the flow capacity of the road network. In contrast, during free-flow states, increasing the number of vehicles increases travel production. However most of the work does not consider the effect of traffic fluctuation in driving local congestion that affect the mobility time cost of the single individuals. For intersecting roads, however, the system resents a lot of fluctuations. This phenomenon poses serious problem to experimental approaches. The ideal of understanding when the congestion is forming and how is propagating is still an open problem and is very dependent on the city [83] and on the heterogeneity of the underlying road network. To experimentally measure such quantities, therefore, one needs big datasets with a small granularity. For example, people use GPS data [24] or loop detectors [108] [109]. In this thesis, we study a GPS anonimized dataset for the city of Bologna (MDT data set [105],[87],[88]). Such dataset is provided by TIM, one of the leader telecommunication companies in Italy with $\simeq 30\%$ of penetration in the mobile phone population. This kind of data has been used extensively in the literature ([47],[45],[35],[5],[125]) to reconstruct statistical properties of the population. Lead by the observation that different roads of the city host different velocities naturally without traffic (i.e. highways high speed and urban lower speed) [24] showed that it is possible to generate homogeneous sub-networks according to the traffic dynamics expressed on them [21]. Such an algorithm is therefore useful to light up the presence of multilayer network from trajectory informations, levelling out traffic conditions. The presence of such structure allows for a mesoscopic description of mobility in terms of transition from a layer to the other as in [103], where it is shown that the longer the trip, the higher the probability to use high speed networks. In this context, it has been already shown that within the partitioning given by Fuzzy-c mean clustering, time and length distributions for trajectories in each class are exponential, while aggregated they get similar to power laws. We show that in the case of Bologna, where we measure a characteristic length $\sim 3.8km$ (fig. 4.14), and where we measure high degree of congestion (fig. 4.18), power laws arise in the path time duration, while, in absence of congestion remains exponential (fig. 4.14). This fact suggests, that instead of criticality the power laws can be indication of individual behavior to perform their OD mobility, that uses the different road sub-network according to the distance of the destination. [44] (in this case the multilayer structure of the network). This explanation in particular is explained by MAXENTR principle with an average time and length of percorrence [19], or equivalently, with the hypothesis that people travel, on

average, until they have consumed some energy associated to the time of the travel [72] and similar across modalities of travel [89]. In this thesis we build on these lines and observe for our dataset (including also holidays) the observations just mentioned. Exploiting information about functional subnetwork extracted via c-fuzzy algorithm [24], the goal is to find traffic features for the different homogeneous subnetworks. Since these network are homogeneous in speed, they do not allow the possibility to look for congestion in data. By construction, people that are in the road network when congested are classified to slower classes. This fact hides them (congested paths) from sight if one consider Macroscopic Fundamental Diagram, or speed averages. We propose a method to reassign paths to different classes with the goal of capturing traffic signals (in both velocity and MFD). We see that the method works consistently for all the days analyzed and allows to define a traffic index that captures the percentile deviation of fluxes generation in each sub-network. We conclude therefore that this method of analysis of data is useful to fill the gap about the functionality of the road network and it is effective in the characterization of traffic. We organize the work such that in section 2.2.1 we describe the dataset used, in section 3.2 we outline the analysis done, each of the subsection is the explanation of the methods used. In section 4.2 we outline the results obtained. To recap in a concise way the goal of the study is to highlight the multiscale nature of the urban road network and to characterize both its usage and the statistical laws governing paths. The study will also examine whether congestion effects emerge at the macroscopic level through the fundamental diagram-like measures, which is widely used to analyze congestion in road networks. A new index to measure congestion that depends on the multiscale structure is proposed.

1.4 Traffic phase transition

To explore possible scenarios in traffic systems, researchers have developed increasingly complex models capable to simulate Lagrangian microscopic dynamics of such systems [95],[69],[71][121]. The possibility to reproduce huge amount of data does not come however with the understanding of important cause-effect mechanisms that can drive policies. For this task we need parsimonious models that should incorporate a minimal set of parameters and mechanisms [15]. For this reason it is important to create mesoscopic models that are able to propose metrics to characterize collective systems [52],[53]. The frameworks and the metrics provided are varied and are proposed to describe different features of this problem, however they share the fruitful lens of statistical mechanics approach to such problems [56]. This work aims to develop a macroscopic understanding of urban traffic dynamics and their susceptibility to collapse, as a function of car volumes, road network capacity, and individual travel patterns, in relation to the city’s origin-destination (OD) mobility demand—ultimately connected to its urban morphology. When we speak about “morphology” we mean levels of monocentricity. To quantify such a concept usually people use population density [40]. A widespread index was introduced in [102] and it is called Urban Centrality Index (UCI). This index takes into account both difference in density and the spatial distribution. We know [79] that when cities grow, they tend, for cost related reasons, to transition from monocentric to policentric. This transition means an evolution of fluxes and therefore evolution of occupancy of roads [111],[14]. In recent works [117],[74] showed that there exist percolation like transition that can single out bottlenecks of speed in the road network. In this framework the transition can be understood in terms of SIR models where the infection state consists in congestion and healthy to free flow [7] and can be related to macroscopic fundamental diagram. However in our work we care about another one observed in [99]. In this article the authors show in the context of [13],[36] that there exists a directed percolation like phase transition in the fraction people trapped in the road network ($n(t)$) [49] with respect to the loading rate (number of people inserted for unit time in the dynamics) of it (R). In particular, they show for different cities that there is a curious

coincidence between monocentricity levels [120] and how people from a control group remain trapped in the road network in the case of one-lane cellular automata microscopic simulator. In particular, the critical exponent in the phase transition, is close to the universality class of 1D directed percolation in the case of monocentric cities and 2D for polycentric [56]. In this study case we focus our study to the cities of Boston, Rio De Janeiro, Lisbon, San Francisco and Los Angeles building validated origin destination [83] and running simulations for Rio and San Francisco. The focus of this work is therefore to develop a pipeline in which we change population distribution [120] so that we reproduce different levels of UCI for each different city. For each of such configurations we explore the phase transition [49] by launching microscopic multi-lane microscopic vehicular simulations of traffic [121] for different loading rates. We therefore focus our attention on the evolution of the universality exponent with respect to the UCI. In chapter 2.3.1 we describe the data we work with, the different sections of chapter 3.3 we introduce the mesoscopic model used for the study of the phase transition 3.3.1 what we mean about monocentrism/polycentrism 3.3.3, how do we transition from one to the other 3.3.6 and what is meant by the universality class directed percolation 3.3.4. In 4.3 we describe the results. The results are shown for the case of San Francisco and Rio De Janeiro where it is visible that the observed transitions cannot be associated to the universality classes of 1-2D directed percolation.

1.5 Venice

Venice is located in the northeastern region of Italy and comprises both a modern, industrialized area—accessible by various modes of transportation—and a historical center, which is effectively an island connected to the mainland via a narrow strip of land. The historical center constitutes the primary focus of this study, as it represents a paradigmatic case for urban analysis. Constructed on a shallow lagoon, the city was founded by embedding wooden piles closely together into the mud until reaching stable ground. Over the centuries, these piles have compacted and mineralized, forming a remarkably durable foundation that has preserved the city’s structural integrity.

This distinctive foundation has profoundly influenced the urban morphology of Venice, resulting in a predominantly pedestrian environment interlaced with canals and connected by numerous bridges. Access to the historical center is limited and can occur via train or car over the causeway, or by waterborne public transportation (e.g., vaporetto). Due to its architectural uniqueness and wealth of historical monuments, Venice ranks among the most visited cities in Italy, drawing substantial numbers of tourists throughout the year. According to the annual report by the Department of Tourism, ~ 10 million visitors were recorded as “arrived” in the historical center alone in 2019, approximately

86 % of whom were international tourists [60].

The historical center covers a surface area of 6.7km^2 and hosts a resident population of approximately 55,000 individuals. During major tourist events, this number can temporarily double, exerting significant pressure on the city’s infrastructure and heritage. The coexistence of a fragile cultural landscape with intense visitor flows poses notable challenges for urban governance, particularly in the domain of mobility. The preservation of the built environment is incompatible with excessive tourist presence, which also disrupts the daily lives of residents.

Nevertheless, the substantial economic benefits derived from tourism make restrictive policies—such as strict limitations on visitor numbers—politically and economically contentious. Therefore, our study aims to analyze and estimate mobility patterns within the historical center, focusing in particular to touristic flows. Such flows (see 2.4.1), are collected via 53 cameras in different positions in Venice. The reconstruction of such flows is then simulated via agent based model (see 3.4) for pedestrian mobility.

Chapter 2

Data

In this chapter, we begin with a brief overview of the datasets commonly used in the literature on human mobility. We then introduce the specific datasets employed in the various studies presented throughout this thesis.

Each dataset described here corresponds to at least one dedicated section in Chapters 3, 4, and 5. For completeness, we list various data sources even if they are not directly used in our work. For a more comprehensive review, the reader is referred to the Data Sources section in [12]. The study of human mobility has undergone a major shift with the advent of complexity science, which provides a robust framework to explore emergent patterns, feedback mechanisms, and structural organization in cities. While early mobility models often assumed equilibrium distributions and relied on static infrastructure maps or small-scale surveys, the tools of complexity science allow researchers to understand cities as dynamic, out-of-equilibrium systems driven by interacting components—individuals, institutions, infrastructures—whose aggregated behavior generates highly structured but evolving mobility patterns.

Complex systems approaches have enabled a better understanding of urban infrastructure as spatial networks, where topology and geometry jointly determine the structure and evolution of transportation systems. Cities’ street and road networks, for example, can be described as planar graphs whose properties—such as the distribution of block sizes, street lengths, and betweenness centrality—exhibit universal scaling laws and non-trivial hierarchies [16, 80, 84, 27]. These insights have led to a redefinition of key urban metrics: rather than analyzing road capacity or density in isolation, researchers now quantify the emergence of hotspots, spatial hierarchies, and scale-free distributions in mobility networks, reflecting the system’s far-from-equilibrium organization [82, 17].

The increasing availability of granular, real-time data has played a pivotal role in this transition. Mobile phone data, in particular, allow researchers to reconstruct origin-destination (OD) matrices with unprecedented resolution, enabling the dynamic observation of commuting flows and the identification of structural regularities across cities [26, 78]. Using such data, mobility networks can be reduced to coarse-grained fingerprints that retain essential information about spatial organization and flow hierarchy, revealing for example how the proportion of “random” commuting flows increases with city size, and how the system self-organizes into a polycentric or fragmented structure as population grows [78].

Beyond these descriptive advances, complexity science provides a modeling toolkit—through gravity laws, entropy-maximizing approaches, or agent-based simulations—that captures the interplay between infrastructure and behavior. In particular, models like the radiation model or closest opportunity models use minimal assumptions to predict how people navigate the trade-off between distance, opportunity, and resource constraints [110]. These models shift focus from equilibrium trip distributions to mechanisms of search and allocation, highlighting the adaptive behavior of agents in non-stationary urban environments.

Altogether, complexity science has transformed the study of human mobility into a data-rich, mechanism-oriented, and multi-scale field. It bridges the structural properties of spatial infrastructure [33, 16] with real-time behavioral dynamics revealed by large-scale data, enabling researchers to analyze cities as evolving, out-of-equilibrium systems shaped by continuous feedback between flows, form, and function.

The spatial and temporal resolution of data used to analyze human mobility has significantly improved over time, driven by advancements in technology.

One of the earliest approaches to studying human mobility involved estimating aggregate population flows using census data. These data are typically collected every ten years through national surveys, in which householders report demographic and socio-economic information about their family members. Census-based datasets are well suited for analyzing low-frequency mobility patterns, such as commuting or internal migration, especially when combined with population estimates like those from [1]. Such analyses are often modeled using gravity-like frameworks [12].

While census data provide a macroscopic view of mobility, they are inadequate for capturing daily movement patterns. One of the first attempts to study higher-resolution spatial dynamics was through tracking the physical movement of money, as in [22]. However, more granular insights were soon enabled by mobile phone data, particularly Call Detail Records (CDRs) [47]. An even better source of data in this direction is represented by XDR. These data are timestamped and georeferenced records of network mobile data sessions, generated by each mobile device interacting with the operator’s network and can be used in many different ways of modelling (see [65]).

CDRs are collected whenever users interact with a cellular network (e.g., making a call or sending a message). These records provide time-stamped locations associated with cell towers, which vary in spatial coverage. Although the resolution is too coarse for reconstructing trajectories at the level of road networks—making them unsuitable for traffic modeling—CDRs have proven valuable for studying general patterns in human behavior. A known limitation is their inherent bias, as they only capture users when they are actively using their phones [47].

Recent technological advancements have enabled the collection of precise spatiotemporal trajectories through GPS data. This has allowed researchers to study both behavioral and traffic-related aspects of mobility simultaneously [81]. The widespread use of smartphones and GPS-enabled devices has resulted in an explosion of available data on human movement, ranging from individual-level dynamics to large-scale population patterns [104]. These datasets support analyses at temporal scales from minutes to decades and spatial scales from individual users to entire nations.

Online services such as Facebook, Foursquare, and Flickr also contribute valuable geo-tagged data. These are generated whenever users enable location tracking for content—such as check-ins, photos, or posts. The resulting datasets include geographic coordinates, timestamps, and additional metadata (e.g., place names or user interactions). These data have been used to study jump-length distributions, visitation frequencies, and mobility predictability. They also allow for estimating co-location and potential social interactions between individuals [66], [106].

An additional source of high-resolution data comes from GPS devices installed in vehicles [19]. While such data can be affected by satellite signal quality and positional accuracy, they offer high temporal resolution and precise measures of velocity and distance. These data are instrumental for studying distributions of path lengths, i.e., the distances traveled between engine start and stop events and congestion.

From an Eulerian perspective, traffic dynamics are often studied using sensor data collected from infrastructure deployed along roads. These sensors measure vehicle speed and density over short time intervals and specific road segments. Such datasets are commonly used to investigate fundamental diagrams of traffic flow, shockwave propagation, and phase transitions in traffic

models, including those within the three-phase traffic theory framework [34], [118], [68].

2.1 Data for Mobility changes during Covid-19

2.1.1 Data

The dataset introduced in this section comes from Cuebiq. Cuebiq Inc. provided the human mobility data used here for research purposes. This data was collected from anonymous mobile phone users who have opted-in to give access to their location data anonymously, through a GDPR compliant framework. Researchers queried the mobility data through an auditable, cloud-hosted sandbox environment, receiving aggregate outputs in return. The datasets contain records of people living in California and Massachusetts from January 2021 to the end of September 2021. The datasets are described in Tab.4.1 (California) and in Tab. 4.2 (Massachusetts). We have count 1 milion (181 k) of users in the period preceding lock-down in California (Massachusetts) with a fraction of user in the other people of $\sim 1/2$ in California (~ 0.7 in Massachusetts). In particular, the focus of this study is the analysis of human trajectories. A trajectory, denoted as T_i , is defined as a sequence of stop points, represented as X_i^α . In particular, a stop point is characterized by a set of coordinates, specifically latitude (lat_i^α) and longitude (lon_i^α), corresponding to the central location of a square with an area of 490 square meters. The coordinates are determined based on the location of a mobile phone signal, which must remain within the same grid square for at least 10 minutes. Due to the nature of the dataset—GPS data collected through Cuebic’s location intelligence service, often characterized by short bursts of high-frequency sampling—it was necessary to apply a filtering step to ensure consistent positional tracking. In particular, stop points were discarded if their start time occurred within a defined threshold of the previous stop point’s start time, in order to avoid redundancy and artifacts from temporally clustered data. In this discussion, a threshold of 1 hour (t^{burst}) is used; however, alternative threshold values of 8 and 24 hours were also applied and are available in the Appendix for reference A.1.

2.1.2 Notations

In this section we introduce the objects we are going to use in section 3.1 where we describe the metrics to quantify the changes in mobility and in the results section 4.1.

The summary of the notation in 2.1 that are going to hold throughout these sections.

We call Users the set of people under study we index them with greek letter α .

We call grid the geometry of squared N_{grid} cells defined by geohash 7 tiling with side size $L_{cell} = 70$ m, we index each cell with roman letters i,j,k. We therefore denote a single cell of the grid as g_i .

A stop point (sp_t^α) is a couple of coordinates ($lat_t^\alpha, lon_t^\alpha$) and a starting (t) and ending time ($\hat{t}$) such that the mobile phone’s signal remains within a cell of the grid for a minimum duration of t^{sp} 10 minutes: $\hat{t} - t > t^{sp}$. We consider the time of the stop point as the beginning of the stop. We denote stay time $\delta t_t = \hat{t} - t$ and is variable in a trajectory.

To each user it is associated a trajectory. A trajectory $traj^\alpha$ is a succession of N^α stop points distributed in time $T^\alpha = \{t_\tau : \tau = 0, 1, \dots, N^\alpha\}$ and the time duration of the trajectory is $|T^\alpha| = t_{N^\alpha} - t_0$. The points are either characterized as latitude-longitude pairs (angle) or projected in the tangent space where we can compute distances in meters. We define trajectories in three equivalent ways as

- $traj^\alpha = \{sp_t^\alpha\}_{t \in T^\alpha}$ (succession of lat lon)
- $traj^\alpha = \{g_{it}^\alpha\}_{t \in T^\alpha}$ (succession of cells)

- $traj^\alpha = \{(x_t^\alpha, y_t^\alpha)\}_{t \in T^\alpha}$ (succession of points in the tangent space).

We call $Trajs = \{traj_\alpha\}_{\alpha \in Users}$ the set of trajectories. Each trajectory is characterized with an interval $[t_{start}^\alpha, t_{end}^\alpha]$ that is contained in the overall period of study $[T_{start} = 15/01/2020, T_{end} = 01/09/2020]$.

We further divide the overall period into three different periods representing different stages of lockdown restrictions that we index with suffixes $\{pre, during, post\}$ and generically referred to with suffix p .

This partition is defined by the ordered time intervals $\{[T_{start}^{pre} = 15/01/2020, T_{end}^{pre} = 15/03/2020], [T_{start}^{during} = 15/03/2020, T_{end}^{during} = 15/05/2020], [T_{start}^{post} = 15/05/2020, T_{end}^{post} = 01/09/2020]\}$ and induces a partition in the set of users that we call $\{Users^{pre}, Users^{during}, Users^{post}\}$. We assign a user to the class p that first satisfy $t_{end}^\alpha \leq T_{end}^p$.

In the case user α spans multiple periods, the trajectory is cut into multiple sub-trajectories and each sub-trajectory is considered for the different partition $Users^p$.

We define also a time burst filter t^{burst} as the minimum time interval within two successive stop points. We define as the geodesic distance $d_g(\vec{x}, \vec{y})$ between two points $\vec{x}, \vec{y}$. In the end we define the mobility network of user α as $G^\alpha = (N^\alpha, E^\alpha)$. This is a weighted directed graph where the nodes (n) are the stop points, while the edges (e) are the transitions between successive stop points and the weight w_e is the number of time such transition happened. The graph is naturally loop-less. We call $k_n^{\alpha, in}$ the number of time node n is visited.

$grid, (i,j,k)$	set of cells geohash 7, (indices)
$d_g(\vec{x}, \vec{y})$	geodesic distance between two points $\vec{x}$ and $\vec{y}$
Users	Set of people
$Trajs$	Set of all trajectories
α	index user
l^α	length of the trip user α
T^α	set of times for each stop point for trajectory α
$ T^\alpha $	time duration for trajectory α
T_{start}, T_{end}	start and end time analysis
T_{start}^p, T_{end}^p	start and end time of period p
$t_{start}^\alpha, t_{end}^\alpha$	start and end time user α
τ	integer index time of a stop point in a trajectory
$t^{sp} (=10 \text{ min})$	interval definition stop point
δt_t	duration of sp_t
t^{burst}	minimum time between successive stop points (1 hour in the main text)
$Users^p$	Set of people per period p
$Trajs^p$	Set of all trajectories per period p

Table 2.1: Standard Definitions

2.2 MDT Data on the metropolitan area of Bologna

2.2.1 Data

TIM MDT (mobile data terminal) technology records the GPS positions of mobile devices in an anonymous way ([105],[87],[88]), when they are connected to the telecommunication network. Due to the activities that last over time, these data allow to detect individual paths on large areas. The MDT dataset was made available thanks to a collaboration with TIM, one of the greatest telecommunications companies in Italy that provides service to approximately one third of the mobile device population in Italy. In the thesis we analyze the data collected in the metropolitan area (MA) of Bologna, but we restrict the analysis to the mobility observed in the Bologna urban area (see fig.2.1) to study emergence of congestion effects. The MDT data

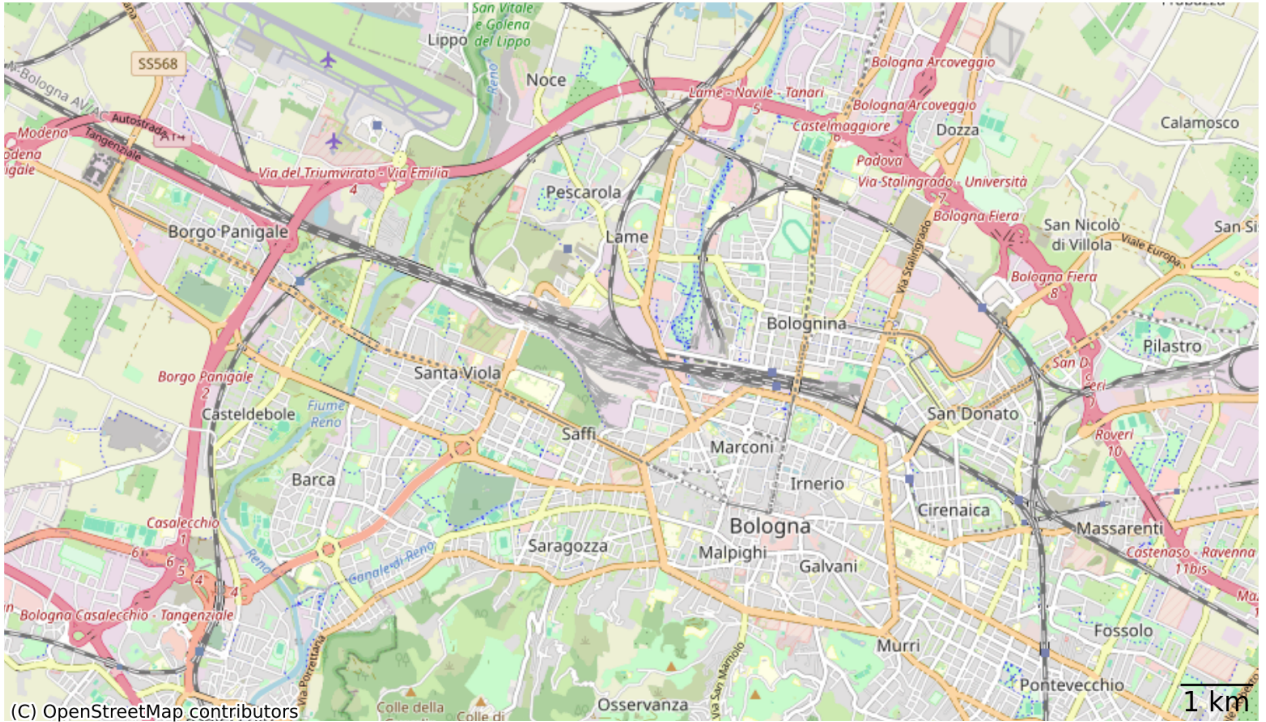


Figure 2.1: Bounding box used for the analysis of the road network of Bologna

are collected during different days in the period [2022 – 12, 2023 – 3] to study different traffic conditions in the area.

The MDT dataset provides information about a sub-sample of the TIM mobile devices connected using the Long-Term Evolution (4G-LTE) wireless broadband communication recording the GPS positions at fixed time interval of 5 seconds during a communication activity. An anonymous id is associated to each activity. The same mobile device may appear under different user IDs across activity records, making it difficult to consistently track a single individual. This means that exists an injective non-invertible function from trajectories to mobile-phones. Each trajectory is therefore associated to a single mobile-phone that however can be associated to multiple trajectories (IDs). Specific algorithms have been developed to georeference the GPS points on the road network with controls on the GPS data quality and to reconstruct the mobility paths [91] using MTD data. In a previous work [89] we have also shown as the MDT paths allow to detect different types of mobility in a urban area (e.g. the cycling and pedestrian mobility) and to classify the urban traffic according to the different use of the road network. We apply these procedures to reconstruct georeferenced mobility paths in the Bologna road network using the available datasets. In fig. 2.2 we show the evolution of the number of

MDT paths reconstructed during each day. To understand the representativeness of the data with respect to the whole traffic of Bologna, we take advantage from the open data set of the Bologna city [58] that provides the traffic flows measured by a magnetic coil system distributed on the road network comparing the hourly traffic flows computed using the MDT paths and the measured traffic flows during 2022. We define the penetration $p(t)$ as the median value of the ratio of the estimated traffic flow F_{road}^{mdt} using the MDT paths and the measured traffic flow F_{road}^{coil} by the magnetic coils

$$p(t) = median_{Roads_{coil}} \left(\frac{F_{road}^{mdt}}{F_{road}^{coil}} \right)$$

where the median is computed on the subset $Roads_{coil}$ of road network for which the open data are available. Each coil is associated to one and only one road, the flows measured in the road are the sum of flows in both direction. In this way we are able to estimate the penetration of the traffic of the MDT path that ranges from 2 to 4 % consistently with the estimates in a previous work[92] (see fig. 2.2).

2.2.2 Notations

We here introduce the objects we are going to use in both methods 3.2 and results 4.2. The name and basic definitions of variables used are summed up in 2.2. In this section we introduce the definitions of the objects and we summarize the notations that we use throughout the thesis in the table 2.2. An MDT path is defined as a sequence of latitude–longitude coordinates recorded at discrete time points t_i , represented as $\mathbf{x}^\alpha = \mathbf{x}^\alpha(t_i)$, where $x^\alpha(t_i)$ denotes the position at time t_i , and the index α uniquely identifies the path. The analysis is performed on a daily time interval so that $t_i \in [02 : 00, 23 : 59]$. In the path reconstruction process we have discarded the GPS data when the distance between successive points is smaller than 50 m and the estimated speed (i.e. the average speed for successive points) is bigger than 150 km/h. We call $Trajs$ the set of all MDT paths. To each path we associate the succession of crossed roads. We call such object $\mathbf{x}_r^\alpha = (\mathbf{x}_{r_t}^\alpha)$ where now the subscript $r \in Roads$ is the identification index of a link of the road network. We call $Trajs_{road} = \{\mathbf{x}_r^\alpha\}$ such a set. In the thesis we study the clustering of the MDT paths into different classes, identified by an integer index $K = \{0, \dots, N_c\}$. The corresponding partition of the sets $Trajs, Trajs_{road}$, will be denoted by $Trajs_k, Trajs_{road}^k$. Such partition, following the clustering procedure in 3.2.5 and the reconstruction algorithm in 3.2.6 produces a classification in the road network as in fig. 3.6 that characterizes the multilayer, functional structure of the road network. We call such partition $Roads_k$. As successively explained, we introduce also $Trajs_k^{new}$ and $Roads_k^{new}$ respectively the partition of the trajectories and roads classified via the hierarchical algorithm presented in 3.2.7. Such sets are important since contain the trajectories that show congestion patterns as measured in fig. 4.18 and the sub-network where that happens in fig. 4.20.

$Users, \alpha$	Set of people in the dataset, index person
T	interval time analysis: $[t_{start}, t_{end}]$
$Trajs$	Set of trajectories
$Trajs_{road}$	Set of trajectories as succession of roads
$Trajs_k$	Set of trajectories that belong to class $k \in K$
$Trajs_k^{new}$	Set of trajectories that belong to class $k \in K$
$Trajs_{road}^k$	Set of trajectories as succession of roads that belong to class $k \in K$
$Trajs_{road}^{k,new}$	Set of trajectories as succession of roads that belong to class $k \in K$ in the hierarchical classification
K	mobility class: $\{k : 0(\text{slowest}), 1, 2, 3(\text{fastest})\}$
$Roads$	Set of roads
$Roads_k$	Sub-networks from projection classes
$Roads_k^{new}$	Sub-networks from projection hierachically filtered classes
δt	15 minutes interval
Δt	1 hour interval

Table 2.2: Standard Definitions

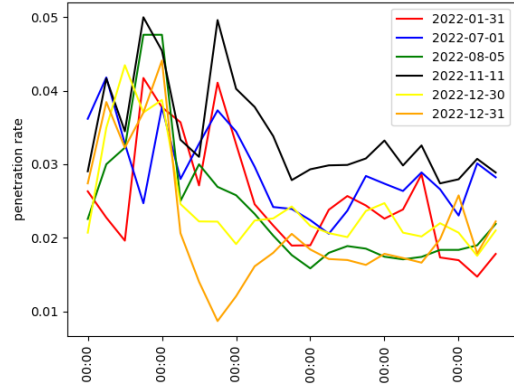
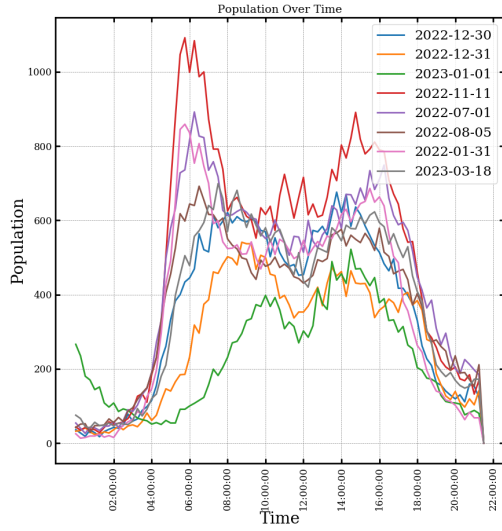
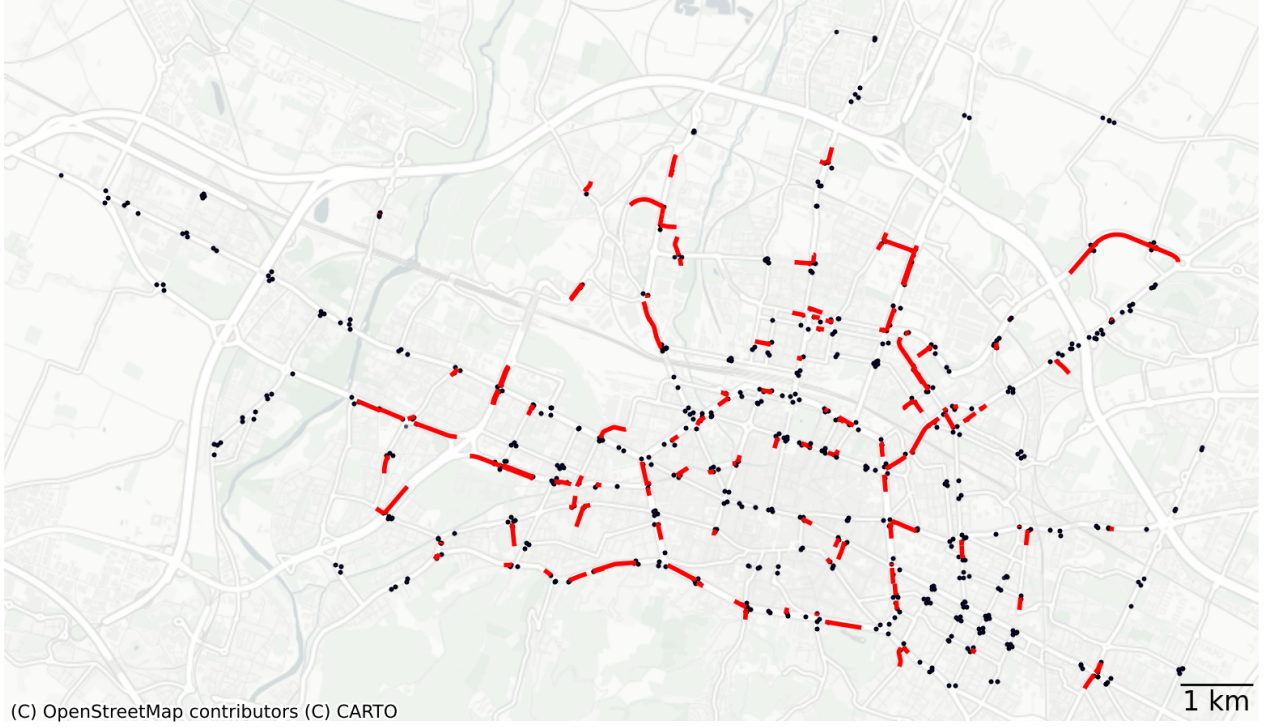


Figure 2.2: Top Figure: The segments highlight the raods equipped by the distribution of the magnetic coils on the Bologna road network. The traffic flows along the red roads define the dataset used to compute the penetration of the MDT paths. Bottom-left picture: Daily evolution of number of MDT paths for the different days in the dataset. Bottom-right picture: the penetration of the data with respect to the baseline given by the coils in the road network.

2.3 Facebook population data for Traffic phase transition

2.3.1 Data

In this section we describe the essential data sources for the study of the phase transition presented in 4.3. We extracted populations from the Facebook dataset that collects information about inhabitants of certain region using hexagonal tiling of a given resolution [1], we extracted this data with a resolution parameter = 8 that corresponds to an hexagonal area of $A_{hexagon} = 0.74km^2$. The data are extracted within the boundary box shown in 2.3 and in the case of Boston, San Francisco, Rio and Lisbon, match those in [99]. The regions for which we have extracted the informations are Boston, Los Angeles, San Francisco, Rio and Lisbon. The bounding box and other informations of interest are reported in tab. 2.3.1. Beside the hexagonal tiling, we produce a square tiling using as parameter the angular opening. For this reason different regions have different size of areas of the single grid depending on the specific projection. The opening angle is 0.02. For boston it corresponds to an area of $1,5km^2$, Rio $3.6km^2$, Los Angeles $2.4km^2$ and San Francisco $2.8km^2$ and Lisbon $4.9km^2$. The mapping of the population from one geometry to the other is such that the fraction of population we assign to a square is the rate of the intersection square-hexagon and hexagon area times the population in the hexagon.

$$P_{square} = \sum_{hexagons} \frac{A_{square} \cap A_{hexagon} P_{hexagon}}{A_{hexagon}} \quad (2.1)$$

city	population (millions)	roads, thousand km	Capacity V, millions	area, 1000 km^2
Los Angeles	10	31.9	265.3	15.6
Boston	4.5	24.8	137.2	5.5
San Francisco	7.4	32.0	171.7	30.6
Rio De Janeiro	12.6	13.8	62.5	9.2
Lisbon	3.8	34.2	70	38.7

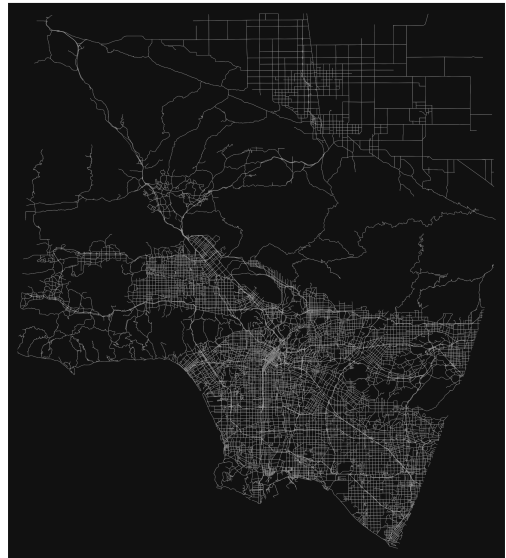
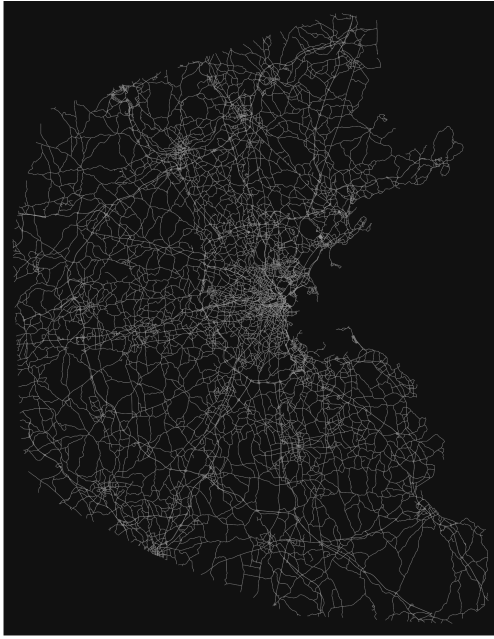


Figure 2.3: Here represented from top left to bottom right: Boston, Rio De Janeiro, Lisbon, San Francisco, and Los Angeles.

2.3.2 Notations

In this section we introduce the objects we are going to work with and the summary of the notation in 2.3 that are going to hold throughout the work. We call the road network $G = (Intersections, Roads)$ that have edges $r \in Roads$ and nodes $n \in Intersection$. We call Users the set of people in the simulation, and $Users^c$ the subset of the control group. We index them with greek letters $\alpha \in Users$ and $\alpha_c \in Users^c$. We call Trajs the set of all trajectories, $Trajs^c$ the trajectories that belong to the control group. A trajectory is a succession of points $\mathbf{x}^\alpha = (\mathbf{x}_t^\alpha)$, where $\mathbf{x}$ stands for (latitude,longitude) coordinates, α is the identification of the person and t is the time. Another complementary notion of trajectory is here used, that is considering the trajectory as a succession of roads. We call such object $\mathbf{x}_r^\alpha = (\mathbf{x}_{r_t}^\alpha)$ where now the subscript $r \in Roads$ is a road intended as a the identification index of a link of the road network. We call $Trajs_{road} = \{\mathbf{x}_r^\alpha, \}$ such a set. We denote t_{start}, t_{end} the starting and ending time of the simulations and t_{start}^c, t_{end}^c the starting and end time of the interval in which the control group enters the road network. We call T the time interval $[t_{start}, t_{end}]$. We define two different time intervals that we call Δt as the time for a single step in the simulation (1 s), δt the time interval of aggregation (10 min) for most macroscopic variables. We call grid the geometry of squared N_{grid} cells of opening angle θ and side size L_{cell} , we label each cell with roman letters i,j,k. We denote origin destination flows at time t the matrix T_{ij}^t , if no time is specified then it refers to the origin destination of the control group. In the end we specify features of the trajectories for user α that are useful for the analysis.

$G = (Intersections, Roads)$	road network
Users, α	Set of people in simulation, (indices)
$Users^c, \alpha_c$	Set of people control group, (indices)
t_{start}, t_{end}	start and end time simulation
t_{start}^c, t_{end}^c	start and end time interval control group
T	interval time analysis: $[t_{start}, t_{end}]$
Δt	1 second, interval update simulation
δt	10 minutes, aggregation interval
$Trajs, (Trajs^c)$	Set of all trajectories, (control group)
$Trajs_{road}, Trajs_{road}^c$	Set of trajectories as succession of roads, (control group)
$grid, (i,j,k)$	set of cells, (indices)
M_i	population in grid i
$T_{ij}^t, (T_{ij})$	flux of people from grid j to i at time t, (control group)
D_{ij}	Distance matrix between centroid of cells i and j
$t_{start}^\alpha, t_{end}^\alpha$	time entrance and exit road network user α
R_α	the ordered set of roads traveled user α
$\langle v \rangle_\alpha$	the average speed of the trip user α
l_α	length of the trip user α
t_α	the duration of the trip user α
t_∞	time defining absorbing state
$t_{n_t=0}$	first time where $n(t) = 0$
$t_{start_{fit}}$	1 hour, time after network unloads when we start to fit
$t_{end_{fit}}$	7 hours, time after network unloads when we start to fit
t_β	7 hours, minimum time to consider the for β fitting procedure
t_0	$t_{start} + t_{start_{fit}} = 9AM$
τ	characteristic recovery time of the network
$N_{ensemble}^{city, UCI, R}$	number of simulations per city, UCI, R = 8
R_c	loading rate transition to criticality
τ_c	recovery time at criticality
Γ	loading fraction
Γ_c	loading fraction at criticality

Table 2.3: Standard Definitions

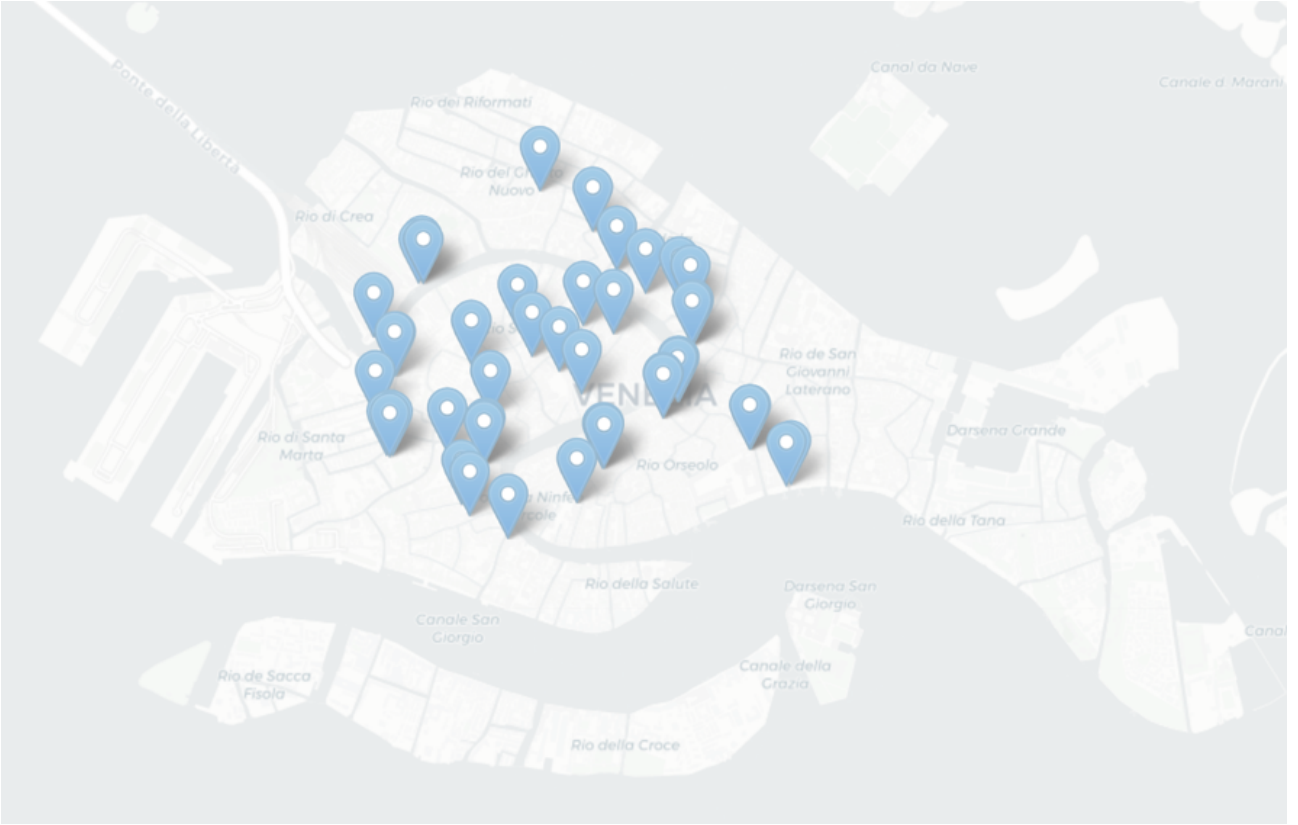


Figure 2.4: Cartography of Venice and position of the 53 flow counting sensors of the Smart Control Room in Venice.

2.4 Pedestrian flow dataset in Venice

2.4.1 Data

The dataset for the city of Venice is the one provided within the project "Venice Smart Control Room" (SCR), in collaboration with the City of Venice, Telecom and Fabbrica Digitale company [62],[63],[61] and consists of flows collected from cameras in 53 different points in Venice. This project aims to create a sort of urban intelligence platform monitoring the entire lagoon city. It consists of a series of data and flows from different subsystems and sensors placed on the territory. Although the SCR collects data coming from different structures, including ACTV/AVM, Centro Maree, and Municipality, we will focus on data provided by TIM (described above) and the network of cameras installed in the critical points of the Venetian mobility. Using Artificial Intelligence and Deep Neural Network algorithms, it is possible to detect people crossing the area covered by cameras and then provide the aggregated count in a specific time interval in real-time. In our case the input data consists of 7 days of flows from 106 different cameras placed on 53 different locations pointing each in different directions. They collect time series of flows updated hourly from 12/07/2021 to 18/07/2021. The aggregation completely preserves privacy rights of people since no further detail about single person is present. These data are going to be used for the comparison with the simulator introduced in 3.4 and to get the results presented in 4.4.

Chapter 3

Methods and Models

This chapter is used as a reference and an introduction for all the methods and models used in result chapter and has been applied to the data introduced in the data chapter.

3.1 Mobility changes during Covid-19

In this section we describe the measures that are used in the case study of 2.1.1. The goal of this section is to define population-aggregated measurements to be computed in every "lock-down related" period to assess the differences in mobility for the different periods. In particular, the goal is to quantify how much lock-down restriction affected mobility at a population level. This is of particular interest since epidemic models rely on equilibrium distributions in mobility in a parametric way [8, 6]. However such behaviors are shown to be perturbed by lock-down restrictions 4.8 both at an aggregated level and disaggregated 4.6. The aim, therefore, is to show experimentally that lockdown produced measurable mutations in mobility behavior at population level. Even though, such results do not inform directly the cited models, this work goes in the direction of providing measures that could be exploited in the derivation of more detailed models, in the direction of the production of synthetic trajectories for meta-populations.

3.1.1 Metrics

We examine a variety of metrics designed to capture different aspects of human mobility, including geometric characteristics, information-theoretic properties, and visitation patterns.

To begin, we use a modified version of the radius of gyration, shown in Eq. 3.1. This quantity measures the average distance from an individual's location to the center of mass of their trajectory at time t , weighted by the duration of each stop. The key modification is that each stop point contributes proportionally to the time spent there—i.e., longer stays contribute more to the radius.

$$R_g^\alpha = \sqrt{\frac{1}{|T^\alpha|} \sum_{t=0}^{t_{N^\alpha}} d_g(\vec{x}_t^\alpha, \vec{x}_{cm,t}^\alpha)^2 \delta t_t} \quad (3.1)$$

Here, $|T^\alpha|$ is the total duration of the trajectory for individual α , δt_t is the time spent at the stop located at $\vec{x}_t^\alpha$, and $\vec{x}_{cm,t}^\alpha$ is the center of mass of the trajectory up to time t . The distance d_g refers to the geodesic distance on Earth.

Another geometric descriptor is the distribution of stop points in an intrinsic reference frame, rescaled by the variances along the principal axes. The computation procedure is described in Algorithm 1.

Algorithm 1 Computation of the stop point distribution in the intrinsic reference frame

for each individual i

$\{ x_i^\alpha \leftarrow \text{latitude}$

$y_i^\alpha \leftarrow \text{longitude}$

$I_{i(xy)} \leftarrow \text{inertia tensor}$

$\text{diag}(I_{i(xy)}) \leftarrow O_i^{-1} I_{i(xy)} O_i$

$(x_i'^\alpha, y_i'^\alpha) \leftarrow O_i(x_i^\alpha, y_i^\alpha)$

$(\sigma_{x'i}, \sigma_{y'i}) \leftarrow \sum_\alpha (x_i'^\alpha - x_{icm}')^2 + (y_i'^\alpha - y_{icm}')^2 \}$

Compute the probability distribution on $\left(\frac{x_i'}{\sigma_{x'i}}, \frac{y_i'}{\sigma_{y'i}} \right)$

We also analyze visitation behavior through two metrics: (i) the cumulative number of newly visited locations over time, and (ii) the average frequency of visits to the ten most frequently visited places per individual (Fig. 4.5). Let $r \in \{1, \dots, 10\}$ denote the rank.

The second metric is computed as:

$$\langle k \rangle^{r,p} = \left\langle \frac{k^{\alpha,r}}{\sum_{r=1}^{\min(10, r_{\max})} k^{\alpha,r}} \right\rangle_{\alpha \in \text{trajs}^p} \quad (3.2)$$

Here, p denotes the time period, r represents the rank of a location in an individual's mobility profile (with $r = 1$ being the most visited), and $k^{\alpha,r}$ is the number of visits made by individual α to the location of rank r . The metric is computed across all individuals active in period p , and captures how visits to top locations are distributed.

To assess the uncertainty in mobility, we compute three entropy measures, each capturing a different level of complexity in the trajectory:

$$H_{\text{random}} = \log_2 n \quad (3.3)$$

$$H_{\text{uncorrelated}} = - \sum_{i=1}^n p_i \log_2 p_i \quad (3.4)$$

$$H_{\text{real}} = \left(\frac{1}{T} \sum_{t=1}^T \Lambda_t \right)^{-1} \log_2 T \quad (3.5)$$

- H_{random} assumes all visited locations are equally likely and corresponds to the maximum entropy. - $H_{\text{uncorrelated}}$ (Shannon entropy) accounts for the frequency of visits to each location, assuming visitations are independent. - H_{real} uses the Lempel-Ziv compression algorithm [126] to approximate the entropy of the actual sequence of visited locations, capturing both frequency and order.

In particular, Lempel-Ziv identifies the shortest substring Λ_t , starting at position t , that does not appear in the preceding sequence. The resulting entropy measure converges to the true entropy of the trajectory as a time series.

These entropy measures provide a progressively refined understanding of trajectory complexity, incorporating increasingly detailed structural information.

Finally, we define the k -radius of gyration, $r_g^{(k)}$, as the radius of gyration computed using only the top- k most visited places. The ratio

$$s_k = \frac{r_g^{(k)}}{r_g}$$

quantifies the influence of recurrent mobility on overall movement. A higher s_k value indicates that an individual’s most frequently visited locations account for a larger share of their mobility footprint.

3.2 Traffic multimodality in Bologna

In this section, we introduce the methods and concepts used for analyzing the data presented in Section 2.2.1, and for producing the results shown in Section 4.2.

The aim of this study is to define an aggregated metric for congestion in the Bologna road network exploiting multimodal structure of the road network extracted from trajectory data collected via GPS from mobile phones. This objective is motivated by a smart-city perspective, where it is crucial to quantify the congestion state of the road network using available data sources.

To this end, we first extract different classes of mobility from the GPS trajectories using the fuzzy clustering algorithm described in Section 3.2.5. These mobility classes are characterized by homogeneous average speeds, which follow approximately Gaussian distributions ordered by increasing mean speed, as shown in Fig. 3.4.

From the resulting partition $Trajs_k$, we extract the corresponding sub-networks $Roads_k$ that support each class of mobility. This structure, constructed according to the procedure in Section 3.2.6 and illustrated in Fig. 3.6, defines what we refer to as the multilayer structure of the road network.

By construction, the sets $Trajs_k$ contain both free-flow and congested trajectories. As a result, congestion effects are not immediately visible in the initial clustering. To address this limitation, we develop an additional method (detailed in Section 3.2.7) to separate congested trajectories from uncongested ones, and to reassign congested trajectories to higher-speed classes.

In this way we have defined classes of mobility in the trajectory space $Trajs_k$, that as a consequence define different sub-networks $Roads_k$ of homogeneous speed (layers) that contain homogeneous speed and we define a mixing procedure that allows to define congested sub-networks $Roads_k^{new}$ with congestion patterns.

This separation allows us to define and compute a meaningful congestion metric, as will be shown in Fig. 4.18.

In the following subsections, we introduce all the necessary concepts and tools required to accomplish this task.

3.2.1 Constructing road network

The first step for any data analysis involving the city is to reconstruct the road network, which means identifying roads and crossroads with lines and points. The best approach is to simplify the complexity of urban space with an essential graph. In this way, any street that connects two crossroads is a link and any road intersection is called node. This allows us to use some concepts from the graph theory like the best path or the degree of nodes and, at the same time, to synthetically keep the information about how users can move on the city graph. The source of cartography information has been Open Street Map (OSM), which is a collaborative project to create free content world maps. Therefore, OSM is considered open data and any user can download maps for whatever purpose. Starting from OSM, different simplifications to the road network are added and here listed (for extensive description see [90]):

- Removal of crossing of degree two: each node (in OSM classification) that has degree two, is changed to point of a road, therefore, just crossings are nodes in our cartography

- removal of short roads: cut those edges that are shorter then $L_{threshold} = 10m$
- Merge sub-graph: compute the largest component and then for each node looks for other nodes that are closer then $D_{threshold} = 60m$, if so, it connects them. This part of the algorithm is essential for not breaking the shortest path algorithm.

3.2.2 Reconstruction of Trajectories with road network granularity

A big part of the job in the reconstruction of the statistical features of mobility at road network scale, is the preprocessing algorithm to standardized the network and the data, trying to avoid noise that naturally is present in the dataset used. Part of my phd time, was used to understand and maintain the code written in [90], I here highlight the salient features.

3.2.3 Filtering Trajectories

The dataset as explained in Chapter 2, Section 2.2.1 is composed by trajectories specified by id,lat,lon,timestamp corresponds to a collection of records with coordinates and timestamp. We apply the following filtering procedure on trajectories to correct for errors of GPS, to specify area of interest and for speeding up the algorithm since some points can be sampled with too much precision (too frequently) without losing informations about movement (chapter 2 [90]). The procedure depends on two parameters

- d_{min} : minimum distance required to generate movement
- v_{inst}^{max} : maximum velocity between two successive points.

This filter allows us to reduce the number of points describing the trajectory and identify and the spatial spike of GPS coordinates by forming clusters of points that are closer then $d_{min} = 50m$ and travelled slower then $v_{inst}^{max} = 150km/h$. It is composed the following steps:

- Generates the first cluster of points with the first point in the trajectory
- Iterates over all points of the trajectory, if the points are closer then d_{min} and slower then v_{inst}^{max} add to the cluster, otherwise generate another one.

At this point the path is the succession of clusters that are centered at their centroid. We define the speed in a path

$$v = \frac{d_{centroid}}{\Delta T_{centroid}} \quad (3.6)$$

where $d_{centroid}$ is the distance between successive centroids, and $\Delta T_{centroid}$ is the distance in time between successive centroids.

The filtering procedure is shown in 3.1.

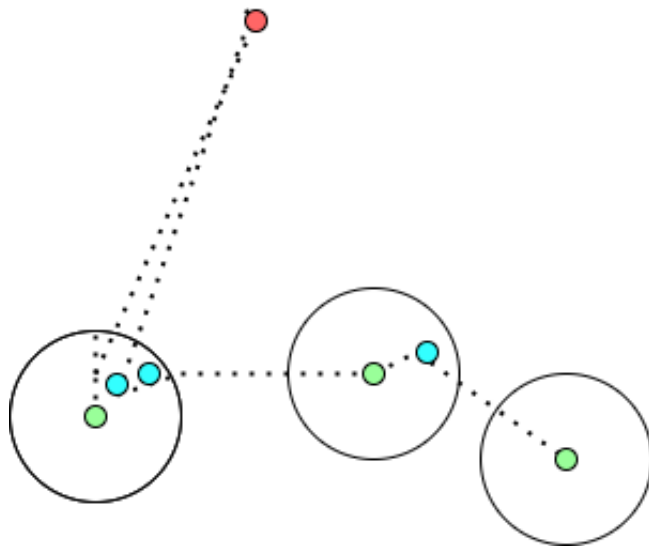


Figure 3.1: Example of spatial downsampling algorithm. Through a loop on records sorted by time, the first is elected as the centroid of the cluster (green point), and the cluster (red circle) continues to accumulate records as long as a record exceeds the distance threshold. At this moment, a new cluster is generated. In this simple example, we pass from 11 records to 4 clusters with a green centroid. These centroids constitute the successive points of a path.[\[90\]](#)

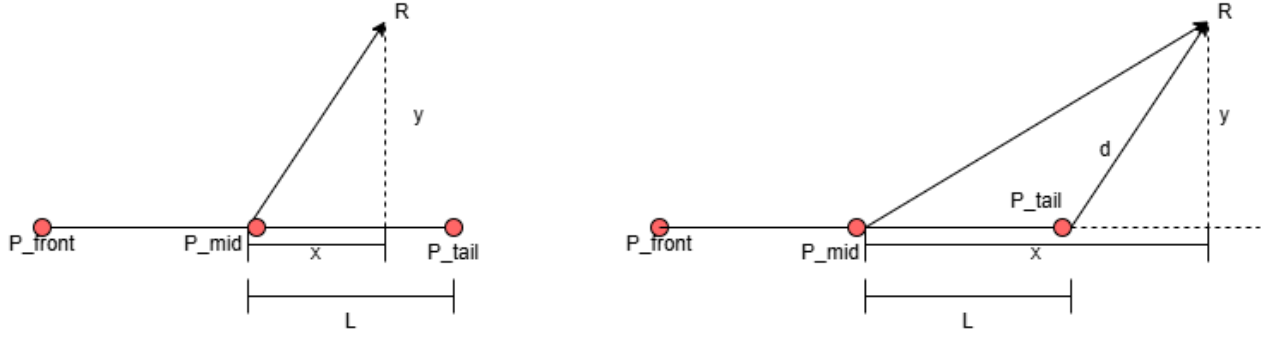


Figure 3.2: Geometry for affinity evaluation of R point with the arc defined by front point P_{front} , tail point P_{tail} and middle point P_{mid} . On the left the intersection with y component lies on the vector defined by P_{mid}, P_{tail} and the distance d corresponds with y. While in the case in which the intersection lies over the vector (right), the distance d is the module of vector P_{tail}, R .

3.2.4 georeferencing and trajectory reconstruction

After reducing points through the downsampling algorithm, it will be necessary to georeference each point on the cartography, reconstruct the covered path between consecutive points and count for each road the number of crossings during the analysis period. By georeferencing, we mean the attribution of a point described by spatial coordinates to an arc, and therefore to a road, of the cartography. A stop point will be georeferenced on the arc, which maximizes a value that we will call affinity, and that is based on the geometrical distance. Before describing the georeferencing algorithm, we need to map the cartography in a grid. This operation will make much quicker all the future passages in which we will need to find the nodes and arches close to a point. Without treating this part in detail, just know that, depending on a settable parameter that we will call $map_{resolution}$, we can obtain $N \times M$ cells with $0 \leq i \leq N$ and $0 \leq j \leq M$ defined like reported in equation 3.7.

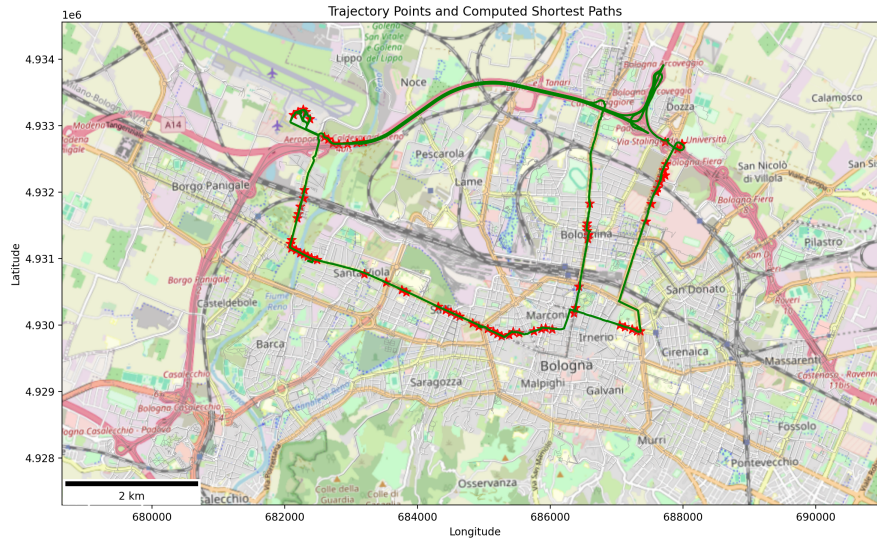
$$\begin{cases} i = (lon - lon_{min}) / (map_{resolution} / ds_{lon}) \\ j = (lat - lat_{min}) / (map_{resolution} / ds_{lat}) \end{cases} \quad (3.7)$$

where $ds_{lat} = 111053.8m$ is the conversion value from degree to meters and $ds_{lon} = ds_{lat} \cos(\frac{(lat_{max} + lat_{min})}{2} \frac{\pi}{180})m$.

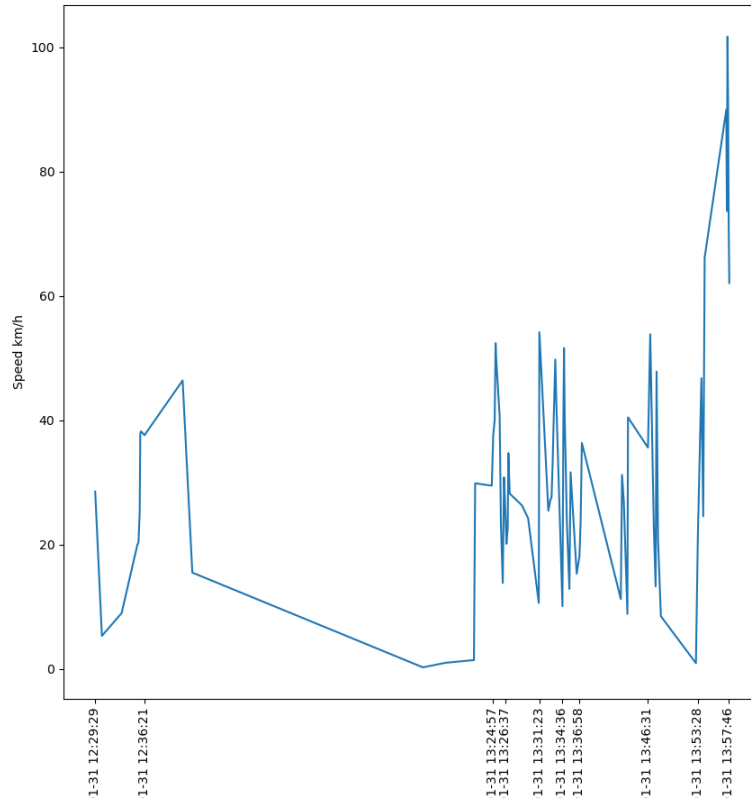
So we generate a map A that associates to each couple (i, j) an object pointing to arches and nodes present in the $cell_{ij}$ and in the nearest neighbors cells. At this point:

- given stop point R that we want to associate to an arc, firstly, we measure the couple (i, j) of the grid
- if $grid_{ij} = \emptyset$ do not assign any road
- else $\forall arc_k \in grid_{ij}$ measure the x and y components from the middle point of arc_k to the stop point R in the reference system (3.2)

The procedure so far is characterized by the bursty nature of mobile phone activities A.11 and subsequent dishomogeneity in density of points time and space. To overcome this limit, we apply shortest algorithm to link successive points in the path. An example of georeferenced path is represented in 3.3, showing both the reconstructed trajectory and the dishomogeneity in time of the signal.



(a)



(b)

Figure 3.3: On the top we can see the reconstructed path on the road network (green) and the centers of the cluster forming the path (red). On the bottom the corresponding speed profile. It can be noted the irregularity of the speed both in the frequency of sampling and in the module. Shortest path reconstruction here allows to fill the gap due to the bias of the signal.

3.2.5 Fuzzy Algorithm: clustering paths

The Fuzzy algorithm, introduced by [21] is a probabilistic clustering algorithm for N_f embedded objects leaving in a space $V \subset R^f$. We used it in the case of Bologna where we extract the sub-network of homogeneous speed 4.2.2. This method requires as input the number of clusters we want to generate (K_f), an acceptance threshold ϵ_f , an exponent m and the matrix $F_{\alpha,i}$, $\alpha \in A = \{0, \dots, N_f\}$ whose rows represent points in V and whose columns represent the coordinates. Each cluster in the algorithm is characterized by a centroid f_i^k with $k \in \{0, \dots, k_f - 1\}$ and i representing the embedding. As a result, $\forall \alpha \in A$ we compute $P(\alpha, k)$ the probability that object α belongs to cluster k .

1. Initialization:

- Randomly initialize the centroids f_i^k
- Compute the probability that each data point $P(\alpha, k) = \frac{1}{\sum_{i=0}^{K_f-1} \left(\frac{\|F_\alpha - f_i^k\|}{\|F_\alpha - f_i^k\|} \right)^{\frac{2}{m-1}}}$

2. Iteration:

- Recalculate the centroid of the cluster as the weighted centroid given the probabilities of membership of all data points O .
- Iterate until convergence or until a user-specified number of iterations has been reached (the iteration may be trapped at some local maxima or minima).

3.2.6 Reconstruction of multilayer character of road network

The method described above has been used to cluster mobility trajectories. It was first introduced in this context in a previous work [24].

The output of the algorithm is a partition of the trajectory set Trajs into clusters Trajs _{k} .

The fuzzy clustering algorithm is applied to N_f trajectories using a feature space of dimension $f = 4$, with a chosen number of clusters $K_f = 4$. The features used, as in the Fuzzy Clustering algorithm [21], are:

- $\langle v \rangle$: average speed,
- $v_{\min}$: minimum speed,
- $v_{\max}$: maximum speed,
- sinuosity = $\frac{d_{sl}}{d_{rn}}$, where d_{sl} is the straight-line distance between start and end points, and d_{rn} is the accumulated distance on the road network.

The feature space is chosen heuristically to capture essential characteristics that distinguish trajectories affected by different levels of congestion and constraints imposed by the road network geometry.

For example, $\langle v \rangle$ strongly depends on both the location and type of road (e.g., highways vs. city center), and whether the trajectory occurred under congestion or free-flow conditions. Similarly, sinuosity tends to be higher in dense urban areas where trajectories involve more turning. Minimum and maximum speeds help to identify fluctuations typical of congestion, such as stop-and-go dynamics or speed oscillations due to traffic lights and spatial constraints.

We chose not to include acceleration due to the bursty nature of the data, which makes higher-order moments unreliable and noisy.

Despite these features, the optimal number of trajectory classes is not obvious. We empirically selected $K = 4$ by inspecting the aggregated speed distributions (Fig. 3.4) and identifying the maximum number of plausible Gaussian modes.

In this simplification, we interpret each cluster as representing a mobility mode constrained by a typical free-flow speed. This assumption is admittedly strong, as congested trajectories may be misclassified into lower-speed classes. We refer to this as the mixing property of the fuzzy algorithm: it tends to group congested and uncongested trajectories based on similar average speeds.

Since the separation between mobility classes is not always clear-cut, a fuzzy classification method is appropriate. It allows for uncertainty in assignment: trajectories that do not exceed a membership threshold of 0.5 for any class are considered unclassifiable. In our dataset, about $\sim 1\%$ of the trajectories fall into this category. The remaining trajectories form the sets Trajs_k .

We index the classes in order of increasing average speed, from 0 (slowest) to 3 (fastest).

With the partition Trajs_k defined, we project this classification onto the road network using georeferenced data (see Section 3.2.4) to define sub-networks Roads_k . These sub-networks constitute what we refer to as the multilayer structure of the road network.

The projection procedure is as follows. For each class k , we:

- Group users by their assigned class;
- Count the number of trajectories traversing each road segment;
- Rank road segments by class-specific flow;
- Select the smallest set of road segments containing a fixed fraction F_k (empirically set to 80%) of the trajectories in class k ;
- If the selected sub-network consists of disconnected components, reconnect them using shortest paths within the full network.

The choice $F_k = 80\%$ follows [24], aiming to reduce noise and ensure a meaningful network topology while preserving major flow routes. As shown in Fig. 3.6, the resulting classification aligns well with known road typologies in Bologna.

This final step addresses the data sparsity problem (i.e., the condition $\frac{|\text{Trajs}_k|}{|\text{Roads}_k||T|} \ll 1$, as shown in Fig. 3.5) by ensuring that each class-specific road layer is a connected sub-network that can support realistic continuous trajectories.

The construction of connected components reflects the assumption that each mobility layer should support coherent urban paths. If the number of additional roads needed to connect components is small, we consider the classification robust and the resulting network a valuable abstraction.

Note that the resulting road layers are not a strict partition of the network: $\text{Roads}_k \cap \text{Roads}_{k'} \neq \emptyset$ for some $k \neq k'$. This overlap again stems from the mixing property of the fuzzy classification, where some roads are used in both congested and free-flow conditions.

Despite this, the classification is remarkably consistent across days: the same sub-networks are repeatedly identified, and the free-flow speed distributions for each class are stable and well-separated see App. 5.1.2. Figure 3.4 shows how average speeds across the four classes form clearly distinct Gaussian distributions with low day-to-day variance.

From inspection of Figs. 4.20 and 3.4, we see:

- Class 0 corresponds to pedestrian and bicycle mobility;
- Class 1 includes small residential and historical center streets;

- Class 2 covers the main arterial roads of Bologna;
- Class 3 corresponds to high-speed roads such as highways.

Since our interest lies in extracting a congestion index, the following Section 3.2.7 introduces a method to further separate congested and uncongested trajectories within the slower classes, resulting in a refined multilayer structure $\text{Roads}_{\text{new}}^k$.

A cautionary note applies to the boundaries of the analysis region. In these areas, classification may be affected by noise, as roads within the bounding box may be connected via shortest paths to roads outside it. This can result in class mixing, as seen in Fig. 4.20, where part of the highway is misclassified as slow mobility in an attempt to connect the periphery of Bologna to the inner city.

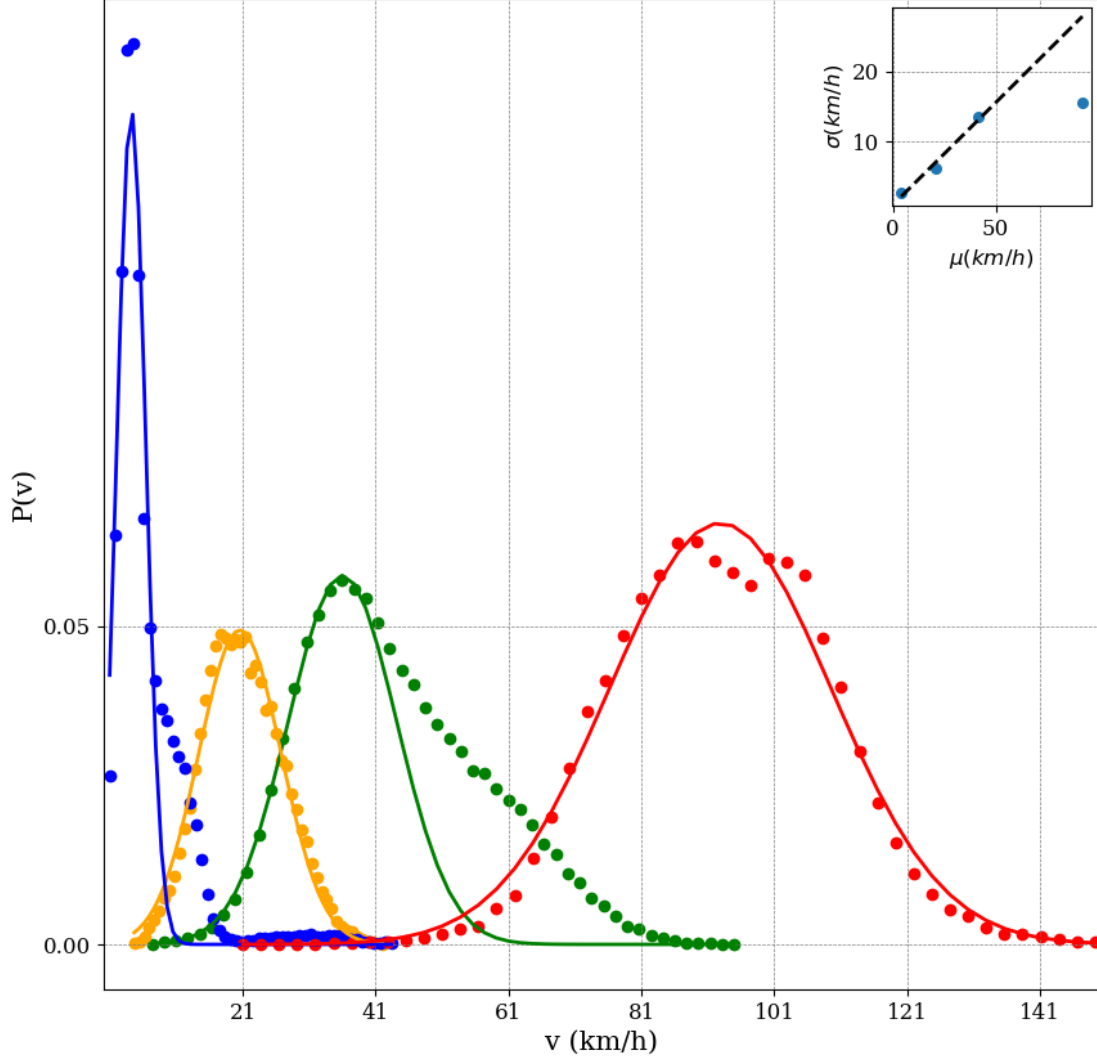


Figure 3.4: The probability distribution of average speeds for the roads on each class: the continuous lines refer to an interpolation with a Gaussian distribution. The averages and the standard deviations for each class are shown on the inset. In the case of class 2 we have fitted the distribution up to 40 km/h, we therefore highlight a deviation from gaussianity. We remark that the same road can be classified into different class according to the traffic conditions and the complexity of mobility paths. This can explain the increase of the standard deviation for the classes 0,1,2 whereas observe the highest speed class 3 deviates from the linear increase since it represents a homogeneous mobility. The results are consistent with the model in [103]. Class 2 comes with a bump in the tail of the distribution, suggesting the presence of a mobility that links the quick urban network with the tangential.

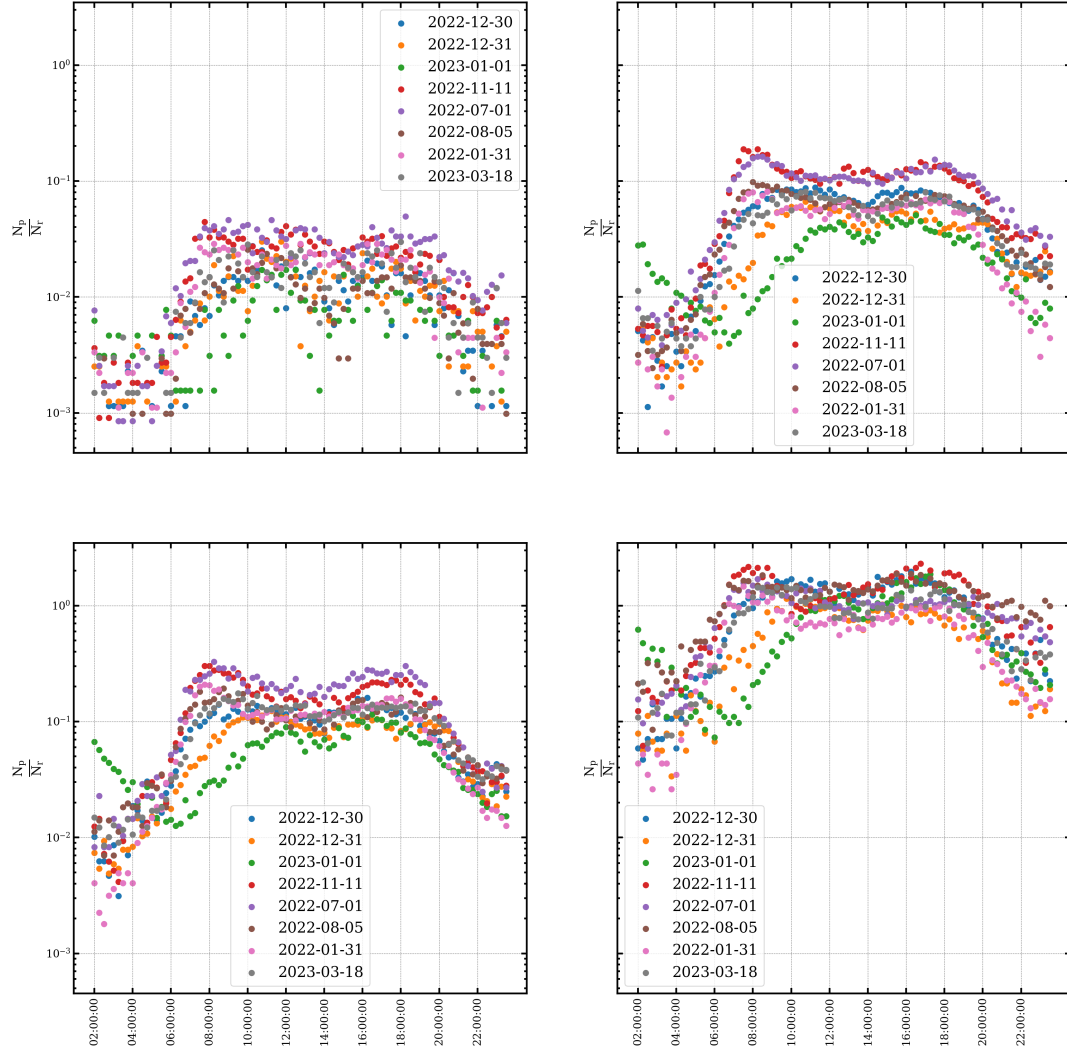


Figure 3.5: Fraction of number of people per road ($\frac{N_p}{N_{tot}}$) for each class of mobility. In increasing order in speed from top left (slowest) to bottom right (quickest) each sub-network. Shown in semi log scale The highway class contains the relative biggest fraction of people, showing the higher capacity of the road network in highway. [4.20](#)

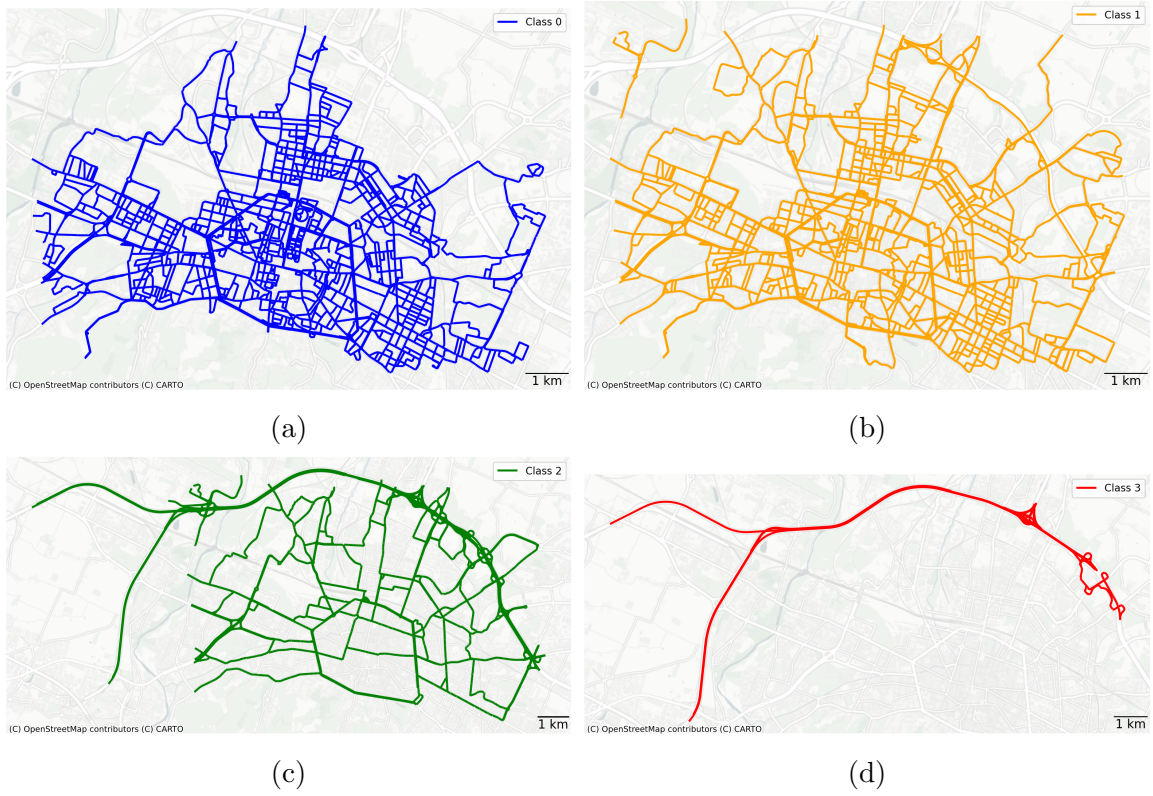


Figure 3.6: Example of multilayer structure induced by the fuzzy algorithm partition of path and successive projection in the road network according to the procedure explained in this section. From top right to bottom left the sub-network associated to the classes in increasing order of speed.

3.2.7 Hierarchical re-assignment class

The road sub-networks $Roads_k$ allow a classification of the mobility paths whose average velocity distribution is interpolated by Gaussian function as shown in Fig. 3.4: i.e. the fluctuations of the average velocity are explained by random effects. However the congested conditions classify paths into classes of lower speed associating them to slower sub-networks.

In particular we observe that the $Roads_0$ seems to contain almost all the roads of the area. To extract information on traffic and congestion effects from the MDT data we have to define non intersecting road sub-networks, that contain roads with homogeneous free flow or geometric features and to reassign the trajectories of the different classes to these sub-networks.

At this aim, we consider the recursive definition

$$\begin{aligned} Roads_k^{new} &= Roads_k \setminus \{\cup_{j=k+1}^3 Roads_j\} \quad k = 0, 1, 2 \\ Roads_3^{new} &= Roads_3 \end{aligned}$$

Our basic assumption is that the free flow features of a road that belongs to unique sub-network $Roads_k$ are homogeneous. Therefore the fastest set $Roads_3$ contains only the highway roads and the new set $Roads_k^{new}$ is obtained selecting only the roads that are not contained by other sub-networks. In this way, one subtracts the roads that also belong to higher speed classes from a given sub-network. We exploit the sub-networks $Roads_k^{new}$ to induce a new partition on trajectories $Trajs_k^{new}$ $k = 0, 1, 2, 3$ by considering the fraction of points of each trajectory that fall into roads of the new partition, and re-assign the trajectory to the sub-network associated to the higher fraction. We refer to this new partition as hierarchical partition. In this way we construct $Trajs_k^{new}$ as the sets that contains the new classification and we assume that it contains information on traffic condition of the related new road sub-network. The road sub-networks obtained thanks to the hierarchical algorithm are shown in Fig. 4.20. We observe that the fastest class 3 is essentially related to the highways present in the considered area. The class 2 contains the ring road and the main country roads that are affected by urban mobility, The sub-network related to the mobility classes 1 highlights the urban road network except some central area closed to traffic and the peripheral residential areas where we have probably a great contribution from a slow mobility. As a matter of fact, the classification is essentially stable for all the days.

Once reconsidered this classification we associate the average speed $\langle v \rangle_k^{new}(t)$ on $Roads_k^{new}$ at different time t . Then, using the hierarchical partition we introduce an index for traffic level for each class being inspired by the fundamental diagram and exploiting the null model for free flow state given by the initial fuzzy algorithm.

$$\Gamma_k(t) = \frac{N_k^h(t)(v_k^f(t) - v_k^h(t))}{N_k^{h,max} v_k^f(t)} \quad (3.8)$$

where $N_k^h(t)$ is the number of paths counted in sub-network k at time t using the hierarchical partition and $v_k^h(t)$ is the average speed of the trajectories in $Trajs_k^{new}$. $v_k^f(t)$ denotes the free average velocity of the sub-network computed using the trajectories in the intersection $Trajs_k \cap Trajs_k^{new}$, where $Trajs_k$ are the trajectories detected by the Fuzzy algorithm. $N_k^{h,max}$ is the maximum number of paths measured in the sub-network k varying t . In Fig. 4.18 we report the average traffic index over the all days. This measure is more appropriate if compared to just speed drops to measure traffic signals. When there is no traffic, the fraction in the number of people contributes to lowering the curve (with respect to just speed), while, if there is traffic, the same fraction is close to 1, not contributing to lowering the peak. This property allows to distinguish better the peaks from the valleys.

3.3 Traffic phase transition

In this section, we introduce the concepts underlying the study whose results are presented in 4.3 and whose input data are described in 2.3.1.

The central goal of this study is to investigate the relationship between a phase transition observed in traffic systems and the morphological properties of those systems [99, 49, 56]. In the following, we clarify what is meant by this connection and how it is formalized.

The motivation for this work stems from the observation in [99], where a calibrated cellular automaton model of single-lane microscopic vehicular dynamics exhibited a particular phase transition [49]. This transition was found to lie close to the universality class of one- and two-dimensional directed percolation transitions into an absorbing state.

Our contribution is to revisit this question using a more realistic, multi-lane microscopic simulator, trying to connect the transition to the morphology and seeing whether changing morphology a change in the universality class holds. We show that the occurrence of the transition, and its possible attribution to a directed percolation universality class, depends strongly on the system's morphology. As a result, the apparent universality reported in the single-lane case does apply to more realistic traffic conditions and not straightforwardly extend to different morphologies.

3.3.1 Traffic model and phase transition

The framework we use in the study of traffic patterns is the one proposed in [13]. The model is defined as follows. Starting from an un-loaded network, we impose a constant input of newly created users. The insertion of travellers starts at $t_{start} = t_{start}^c = 7AM$ and for each simulation time step (Δt), R travellers are inserted in a node of the road network, picked randomly from the origin cell of the pre-computed origin destination (OD) matrix for one hour until t_{end}^c . The people inserted in this interval of time ($[t_{start}^c, t_{end}^c]$) are the control group $Users_c$. After the 1 hour, people are continued to be inserted in the road network according to the OD measured for successive times forming $Users$ (in fig.3.7a represented the unload curve for both the $Users$ and $Users_c$). The routing strategies are such that each traveller chooses the shortest path in time, and recomputes it during simulation according to perceived traffic using calibrated re-route strategies [121]. While the loading rate is small enough, we observe an exponential recovery in the number of people that complete their journey 3.9

$$n(t) = n(t_0) \exp^{-\frac{t-t_0}{\tau}} \quad (3.9)$$

defining therefore a characteristic time of recovery for the network τ .

In the case the number of inserted people per unit time Δt in the simulation time is big enough, queues can start from the starting node (with infinite capacity). Given this microscopic dynamics, we are interested in the delivery properties of the road network and the response to increasing loads. To do so we monitor (order parameter) the number of travellers, from the control group ($Users_c$), that have not reached their destinations at each time step t :

$$n(t) = \frac{\sum_{r \in Roads} N_r(t)}{\sum_{\Delta t \in 1hour} R \Delta t} \quad (3.10)$$

where $\sum_{\Delta t \in 1hour} R \Delta t$ is the total number of people inserted in the road network from the control group. This quantity is qualitatively related to another important quantity; the loading fraction (Γ) of the road network that corresponds to the fraction of maximum capacity that congestion state does not allow to use and has the unit of time:

$$\Gamma = \frac{\sum_r p_r l_r}{\sum_r C_r l_r} \quad (3.11)$$

where the p_r is the number of people that have passed through road r , l_r the length travelled, and C_r the capacity ($\frac{veh}{h}$). So it is the fraction of trip that is not consumed with respect to the non congested regime (stationary state). Thus, it captures the spatial distribution of the loading on the available road infrastructure of the city. In [99] they show that Γ and τ are able to capture 3 different phases that identify three different regions of R . The stationary state, where no macroscopic signs of congestion is observed, corresponds to a free flow phase [36], here Γ increases independently of τ (as seen in fig. 3.7). This state is characterized by exponential decay in time of $n(t)$ [83] and absence of queues at nodes. The jammed state, the state for which:

$$\lim_{t \rightarrow \infty} n(t) \propto t^{-\alpha} \quad (3.12)$$

therefore describing a state in which people remain trapped in the system for long times and queues start to emerge. In this regime, we observe an initial linear relationship between Γ and τ , where τ is defined as the time at which $n(\tau) = e^{-1}$.

This linear increase, then, is substituted by an exponential increase, showing a collapse state of the system. The behavior of this curve suggests the presence of a phase transition to jammed state. In particular, it is a transition to an absorbing state, where the system collapse is measured with the number of people trapped in the road network at time that tends to infinity ($n(R, t_\infty)$) [56, 49].

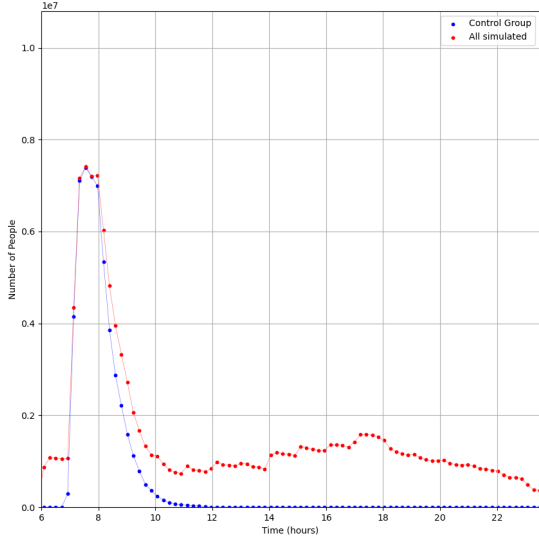
The characterization of such absorbing state is not trivial, since it does not actually exist in traffic systems as re-routing strategies introduce additive noise that are capable to unlock the gridlock. Therefore, our strategy to capture the absorbing state needs to define the time t_∞ that we consider to be the limit to ∞ in 3.12. To do so, we check for each loading rate R whether exponential fit performs better then powerlaw fit (in terms of χ^2). If the exponential performs better, it means that the system is not gridlocked. The first time we get power-law fitting better then exponential, we say that our system is jammed and we mark the respective loading rate as critical (R_c). We transited from exponential unloading to power-law $n(t) \propto t^{-\alpha}$. The α is not universal exponent and depends both on the city and its morphology (UCI). At this point we look at the transition characterized by the power law with exponent β and leading parameter R and measuring at t_∞ the number of trapped people $n(R, t_\infty)$. For each R we define:

$$\epsilon = \frac{R^2 - R_c^2}{R_c^2} \quad (3.13)$$

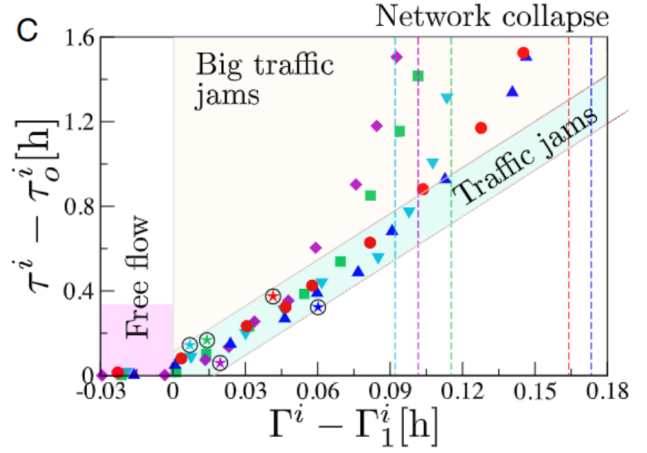
Subsequently, β and t_∞ are chosen by iteratively fitting the power law over all the time intervals $t' = t + k\delta t$ and choosing the smallest error in the fit, that is:

$$n(R, t_\infty) = \min_{t', \beta} \sqrt{\sum_R (n_R(t') - \epsilon^\beta)^2} \quad (3.14)$$

Such a phase transition is the subject of this study whose goal is to see the relationship with directed percolation for different urban forms. As explained in A.3.1 the fit has been computed in log-log scale avoiding therefore to consider $R = R_c$ where $\epsilon = 0$.



(a)



(b)

Figure 3.7: In 3.7b representation of the different phases of traffic as found in [99] in terms of characteristic unloading time τ , in 3.7a the representation of the $Users$ (all) and $Users_c$ (control group)

3.3.2 Traffic Simulator

The traffic simulator used in this study is the one developed in [121]. It implements multi-lane microscopic dynamics consistent with those found in state-of-the-art traffic simulation systems such as SUMO [71]. We highlight this point because the simulator employed in [99] features significantly simpler dynamics and rules, being limited to a single-lane framework.

3.3.3 Monocentrism/Polycentrism

To quantify the amount of monocentrism and policentrism of a system we will make use of a standard measure first introduced in [102], the Urban Centrality Index (UCI). It is defined as:

$$UCI = LCPI \quad (3.15)$$

The first term LC is called the Location Coefficient. It was first introduced by [114] and quantifies the inequality of distribution of either number of jobs or population and ranges from 0 to 1, with 0 completely polycentric, and 1 completely monocentric. In our case we redefine LC keeping the qualitative role of quantifying inequality in the distribution. Inspired by the hot-spots description in [77], we consider the Lorenz curve of the population distribution. The Lorenz curve is a graphical representation of the cumulative distribution function of an empirical sorted probability distribution (fig. 3.8). In our case we take the normalized population distribution to be considered as probability.

$$L(F) = \frac{\sum_i^{F \times N_{grids}} M_i^{ordered}}{\sum_i M_i^{ordered}} \quad (3.16)$$

If all the grids contain the same number of people, then the Lorenz curve would be the diagonal from (0, 0) to (1, 1). In general we observe a concave curve with a more or less strong curvature, and the area between the diagonal and the actual curve is related to the Gini coefficient (the

area within the diagonal and the concave curve), an important indicator of inequality used in economics. In the Lorenz curve, the stronger the curvature the stronger the inequality and, intuitively, the smaller the number of centers. The criterion given for the definition of LC is:

$$LC = F^* \quad (3.17)$$

where F^* is the intersection point between the tangent of $L(F)$ at point $F = 1$ and the horizontal axis $L = 0$. This method is inspired from the classical scale determination for an exponential decay. This approach has the advantage that it does not necessitate prior knowledge about the city (as for example in [46]) and avoids the arbitrariness of defining a threshold in the quantity used to characterize the number of centers. The second quantity is the Proximity index (PI) 3.19.

$$S = \sum_{ij} m_i D_{ij} m_j \quad (3.18)$$

Where m_i, m_j are the fraction of population at grids i and j and D is the distance matrix between the centers of the squares that compose the grid.

$$PI = 1 - \frac{S}{S_{max}} \quad (3.19)$$

The interpretation of PI is the opposite of V , with its theoretical range being $(0, 1)$. Values of PI closer to one mean that population is clustered in one single center (this economic center does not necessarily match the geometric center). If PI is zero, population is as spatially separated as possible. In other words, populations is distributed in a way that maximizes the distances between them. To compute this second index we must highlight that S_{max} is computed by distributing the population in an homogeneous way in the boundary of the geometry considered. In all cases we consider the boundary as the minimal convex polygon that contains the boundary the geometry. In the case of San Francisco, Lisbon and Rio De Janeiro in which we have geographical holes (sea) where people cannot either go or live, we consider as the boundary, the polygon that contains in its inside all the holes. These choices are not optimal but nevertheless well defined way of computing the S_{max} , since all the generated configurations have smaller values of S . As already suggested, the transition between monocentrism and polycentrism is not neat, as they correspond to different regions of UCI. In particular monocentric cities have computed $UCI \simeq 1$ while polycentric $\simeq 0$ values. Our definition of UCI is a little bit different from that found in [102]. It is still composed by two multiplied components, but the definition of them is adjusted to our framework.

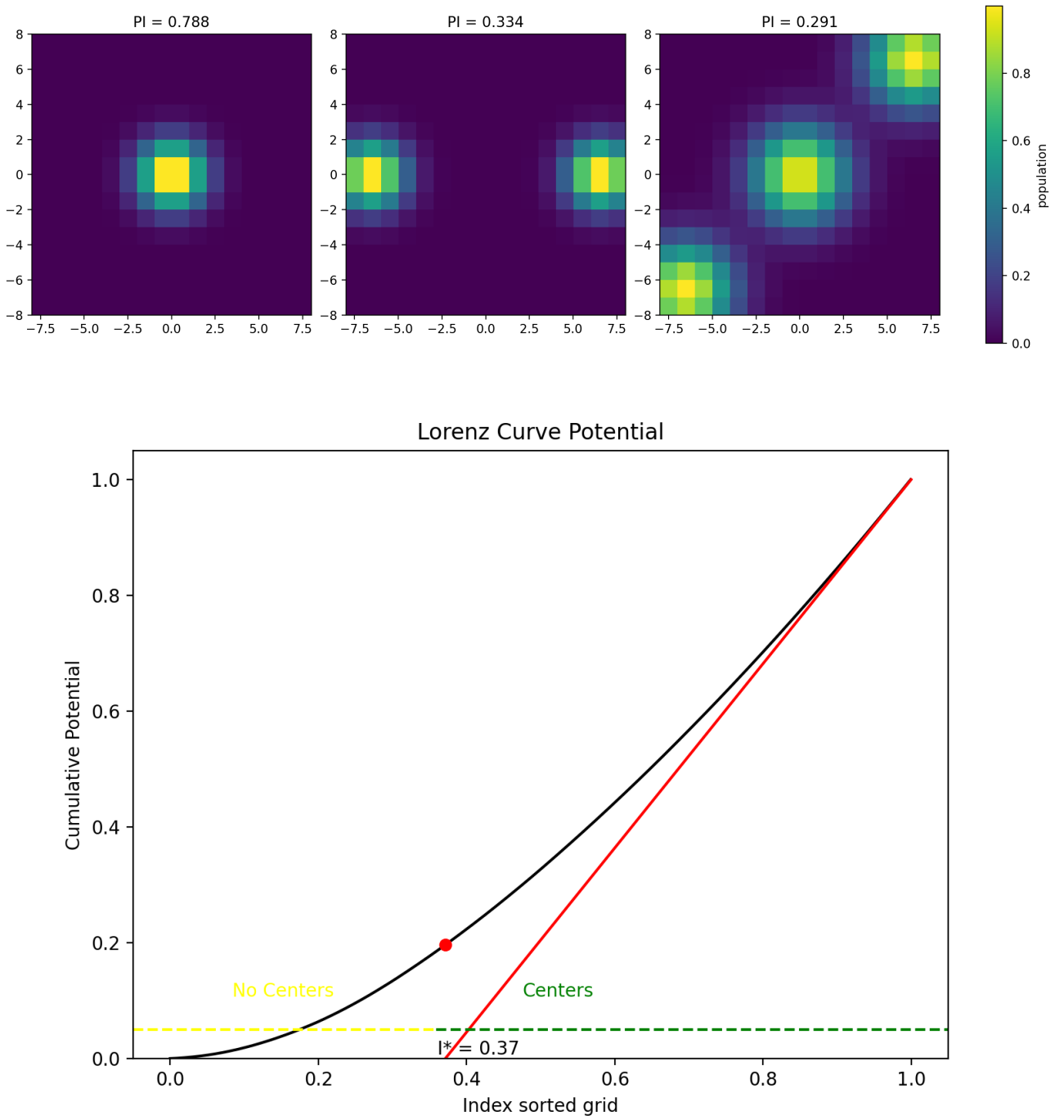


Figure 3.8: On the top the representation of different levels of monocentricity in the spatial configuration (PI). On the bottom the lorenz curve with the schematized principle from [77] to determine which are the cells to be considered centers and which not, determininig F^* as the point in which the tangent to the curve collides with the x axis (LC).

3.3.4 Directed percolation

Directed percolation is the simplest model for out of equilibrium statistical mechanics universality class. The model is usually constructed in lattice $L = (V, E)$ on a time span $T = \{0, 1, 2, \dots, T\}$ where we define on the vertices the quantity $s_{v,t} \in \{0, 1\} \forall t, v \in T, V$. $s_{v,t}$ can be in two states that represent presence or absence of something (initially conceived as wetness or dryness). The model is an update rule of such quantities for all the times and goes as follow: for a given configuration at time t , the next configuration at time $t + 1$ can be determined by extracting $\forall v' \in \text{Neighbors}_v$ a random number $z_{vv'} \in (0, 1)$ (see fig. 3.9). Different microscopic rules produce behaviors that fall into its universality class [56], here, for our purposes we are considering A bond is considered to be open (with probability p) if $z_{vv'} < p$, leading to the update rule:

$$s_v(t + 1) = \begin{cases} 1 & \text{if } s_{v'}(t) = 1 \text{ and } z_{vv'} < p, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, directed bond percolation can be considered as a Markov process with parallel dynamics. As in any dynamic system, we have to specify the initial state. Common initial states are the fully occupied lattice, random initial conditions, and configurations with a single particle at the origin (also called ‘active seed’) [56]. This model generates a universality class that is characterized by the scaling exponents in tab. 3.1

d	β	$\nu_{\perp}$	ν_k	$z = \nu_k / \nu_{\perp}$	$\chi = \beta / \nu_{\perp}$	$\delta = \beta / \nu_k$
1	0.276486(8)	1.096854(4)	1.733847(6)	1.580745(10)	0.252072(12)	0.159464(6)
2	0.5834(30)	0.7333(75)	1.2950(60)	1.7660(16)	0.7955(30)	0.4505(10)

Table 3.1: Critical exponents for directed percolation in $d = 1, 2$.

where β is the characteristic exponent for the density of percolated sites, $\nu_{\perp}(\nu_k)$ for the size of percolated cluster in space (time). The problem we face is to understand whether a few realizations of a dynamic process that is not simulated directly as directed percolation falls in the class of models of directed percolation in the mesoscopic scale. This requires definition of observables that are stable for the phenomenon at hand. The one we look at, is, as already mentioned 3.3.1: β : the exponential for the universality class that describes the phase transition $n(R) = (R - R_c)^{\beta}$

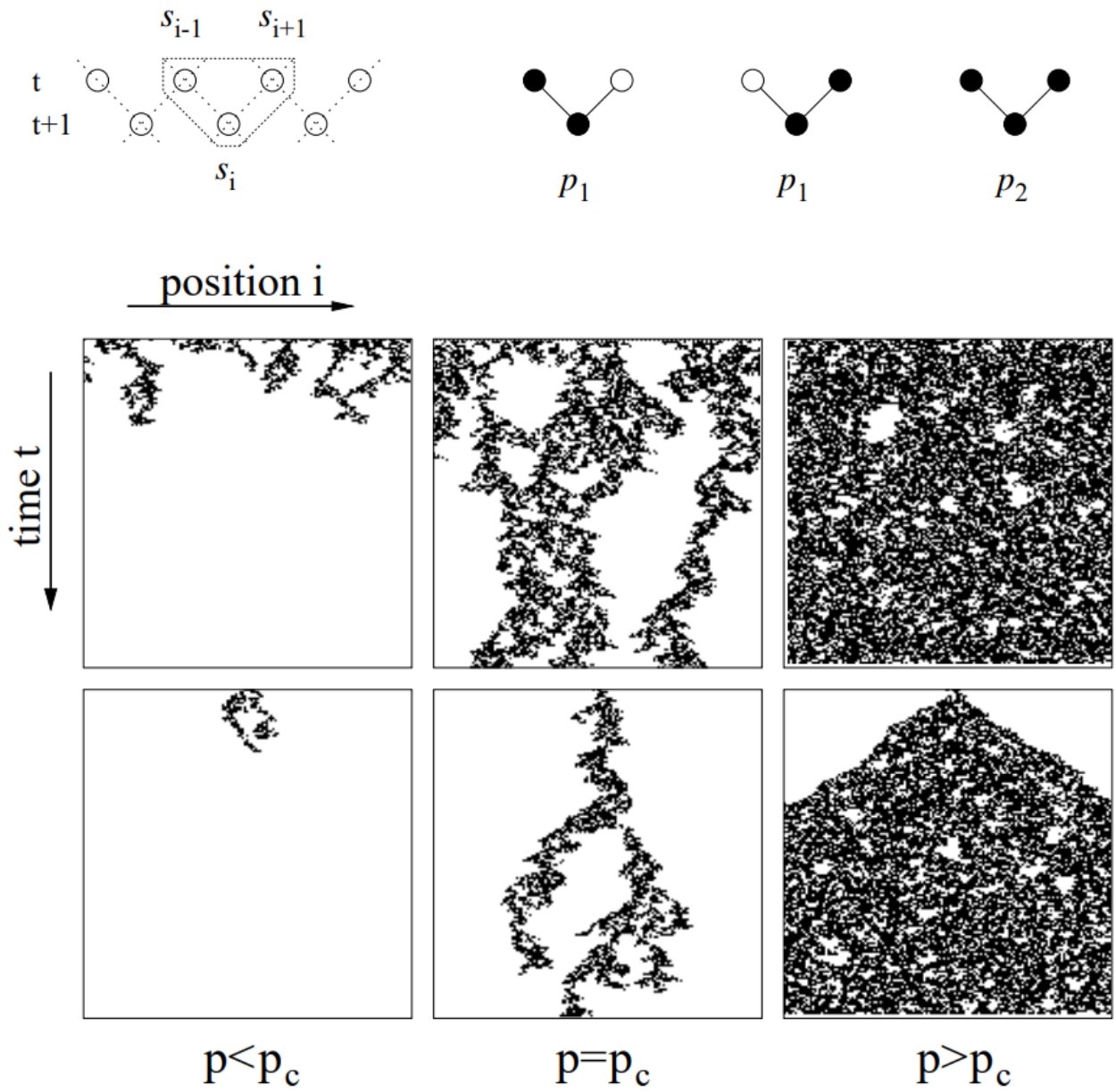


Figure 3.9: On the top the representation of rule associated to the probabilistic evolution of a directed percolation model (Kinzel) in 1 dimensional lattice. On the bottom the representation of the realization of such process at different values of the parameters below, around and above critical value. Both figures are taken from [56].

3.3.5 Gravitational Model for flows

The equation for the flow of commuters between two areas i and j with an exponential deterrence function is

$$T_{ij} = km_i^\alpha m_j^\gamma e^{-\frac{d_{ij}}{d_0}}, \quad (3.20)$$

where k is a constant, m_i, m_j are the populations of origin and destination areas i and j , d_{ij} is the distance between them, and d_0 is a characteristic distance. The model has four parameters to fit (k, α, γ and d_0). For the fitting procedure, we have considered just those couples of cells for which the population was contemporarily different from 0 in the origin and in the destination, and the number of people commuting among the two. We check the validity of the fit by looking at the ratio of flows and the product of the masses, and the ratio of flows and the gravity flows as in [11]. We see that the fit performs reasonably well up to the half of distances available as shown in fig. 3.10 and then deteriorates. We show the fit values in 3.2:

Table 3.2: Parameters gravitational fit for different cities

City	k	α	γ	$\frac{1}{d_0}$
Boston	1.72	0.09	0.13	0.038
San Francisco	0.94	0.14	0.13	0.034
Rio De Janeiro	0.95	0.14	0.15	0.042
Lisbon	1.11	0.13	0.14	0.038
Los Angeles	1.79	0.09	0.05	0.023

3.3.6 Changing the shape of the city: Generation of new flows

The generation of flows depends on three parameters: $N_{centers}$, σ_{center} and d_{center} , where $N_{centers}$ is the number of new centers, σ_{center} is the variance of the gaussian curve used to distribute the population from the center, and d_{center} is the distance from the "chosen center". We define the "chosen center" of the city: the grid where the most people live. At this point we divide the population $N_i = \frac{N_{tot}}{N_{centers}}$ for each new center i . We extract from an exponential distribution in the distance from the "chosen center" (with d_{center} the only parameter) the i -th center. In this process we avoid to consider grids that have less then 300 inhabitants and those that do not contain any road to avoid populated regions that do not have enough infrastructure to handle traffic. Each "new center" now is the center of a gaussian distribution for a population core that redistribute the fraction N_i of overall population. Again, we take care to avoid grids that have no roads so that in the simulations all the people can travel. At this point we apply the gravitational model previously computed to reproduce the new flows T_{ij}^{new} . We show, for different cities the UCI as a function of the LC and PIs generated to transition from monocentric to policentric configuration in 3.11.

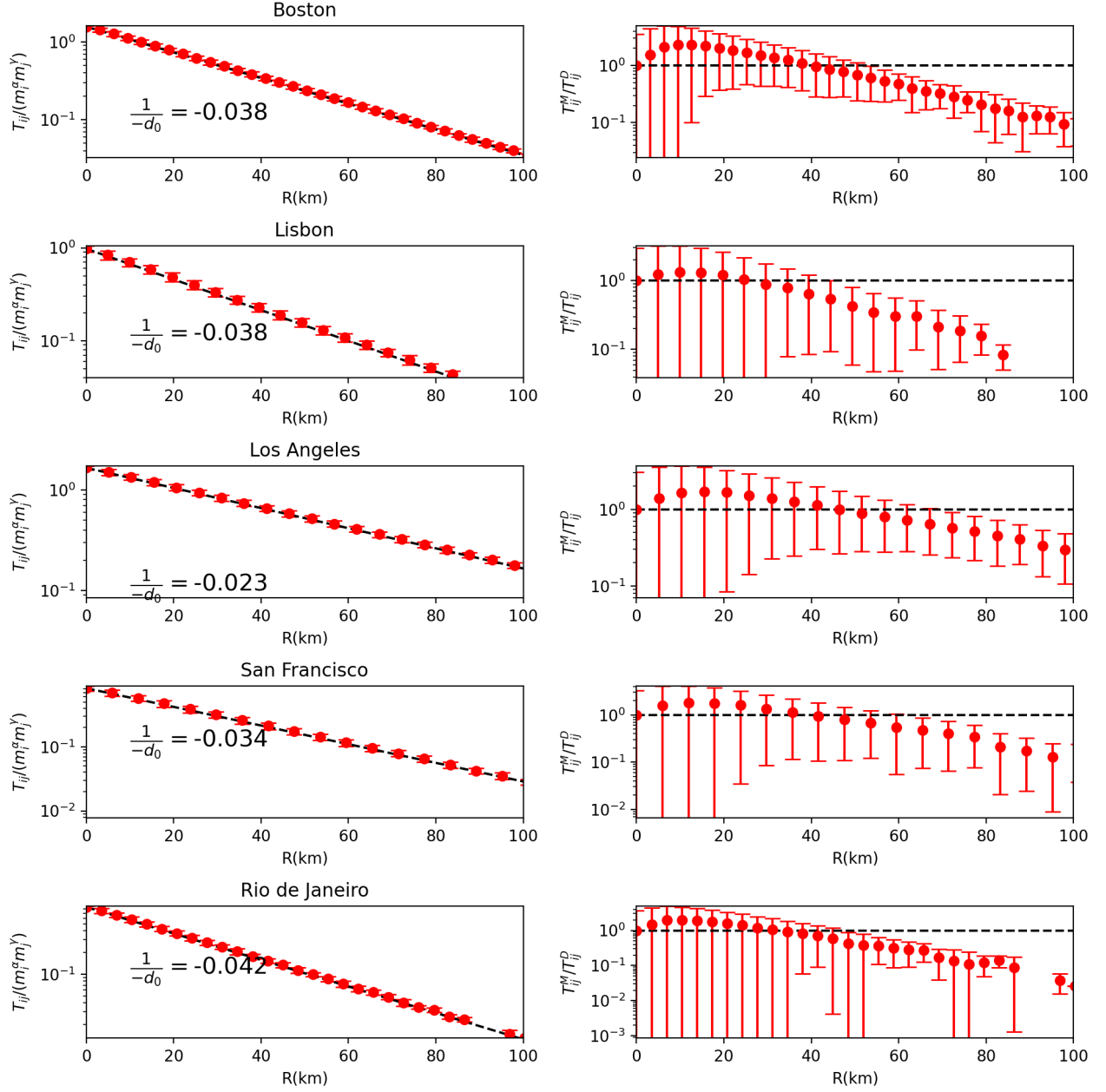


Figure 3.10: On the left the ratio between the flows T_{ij} and the product of fitted masses. It shows the exponential dependence on the distance for each city. On the right the ratio between the flows measured and fitted with the gravitational model. For each city, the fit accuracy deteriorates around half the maximum distance measured. These plots are meant to show the goodness of fit for the gravitational model used.

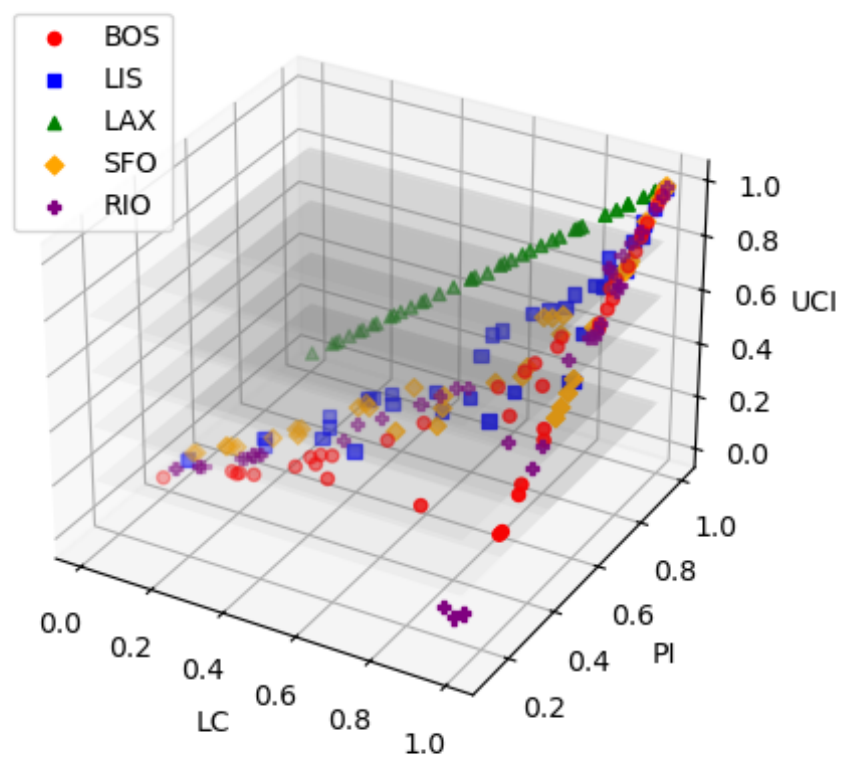


Figure 3.11: Representation of UCIs in terms of LCs and PIs for all the different cities in the study.

3.4 Pedestrian mobility in Venice

In this section, we describe the methods used for the analysis of the data presented in 2.4.1, whose results are discussed in 4.4.

The aim is to validate the performance of a pedestrian simulator for tourist flows in Venice. In the following subsection, we introduce both the simulator and the validation problem. The validation is carried out through a hierarchical analysis of flow correlations across different locations in the city.

Although the contribution of this study is modest in terms of scientific novelty, it provides a useful methodology for calibrating microscopic dynamical models of touristic flows.

3.4.1 Simulator pedestrian mobility in Venice

The simulator of pedestrian mobility in the city of Venice is an agent based model and it is the same of [90]. It has been built to reproduce the flows measured on Venice in 53 different places. The problem that we face, is to validate the goodness of the simulator by comparing the flows produced by the simulator to the experimental flows measured by the cameras. We recall here the principal features and the parameters of interest. cartography is downloaded from OSM and then preprocessed via algorithm in chapter 2 of [90] and briefly presented in subsection 3.2.1. The processing consisted in the simplification of the road network in such a way that a segment is a link if and only if it starts and ends in nodes that are intersection of three of them. To be noted that the simplification of the cartography allows better computational performances. Attractions are points of interests (POIs) of the city, that is, points that attract tourist and attract mobility. They are chosen from historic buildings, squares, exhibitions, museums. They are characterized by their geographical coordinates, attraction level (that can be a vector with any resolution in time default 96 (15 minutes window)), maximum capacity, visiting time table (if available). These object are important to decide where tourists will go. Sources are the object that contain pawns' generation. For visitors and commuters they are usually set in access points of the city bus,train station or port. Pawns that are created in the same place, have the same properties, that will define a set of rules for the dynamics:

- *SRC_DEST* are the pawns with origin and destination fixed. These will be the pawns responsible for tourism since the destinations are set from the set of attractions of the city. These pawns are the ones that simulate touristic flows and are inserted in the simulation using entrance points in the city like the train,bus and Vaporetti station as shown in fig. 4.29
- *SRC_RND* they are generated in a source and assigned a random destination, this is the commuter mobility since they arrive in the city and then are distributed randomly in their work places. This mobility represents commuters in Venice.
- *RND_DEST* the agents are generated in a random point and they have to reach the destination point.
- *RND_RND* agents will be generated in a random point and they have to reach another point extracted randomly. These are meant to simulate residents.

Once defined these objects we need to define rules for insertions (where,when and how many) and decision of trips and roads to accomplish them, and how do they move once they are in the road and how they react to the presence of others. The first task is handled by dividing the time table in intervals of 15 minutes and constraint the number of people entering in the simulation to be the one measured in the cameras at the entry point of the city (this

is done in the case of tourists, the only we do). This procedure corresponds to insert for each SOURCE (red popup in fig. 4.29), $N_{SOURCE,t}$ tourists. For each tourist, the trip is extracted from a probability distribution built in the following way. We choose a minimum (k_{min}) and a maximum (k_{max}) number of places to be visited. Then we build the set of all the possible combinations of successive visits (example [(Ponte Rialto, San Marco), ..., (Papadopoli, Ponte Rialto, San Marco)]). We bin this set by total length (of the shortest path linking successive attractions) and compute the probability of choosing the length via:

$$p(l) = \frac{\beta e^{-\beta l} \frac{1+e^{-\alpha L_c}}{1+e^{-\alpha(l-L_c)}}}{1 + e^{\alpha L_c}} \quad (3.21)$$

Where $L_c = 0.5$ (characteristic length for pawns in distribution of trips), $\alpha = 7.5$ and $\beta = 0.84$ (as calibrated in [90]). Each bin of the distribution therefore contains a set of path of a given length. We extract according the weights of the distribution a bin. The bin now contains a set of path containing, characterized by the presence of $N_{attractions}$ attractions. Each one comes with its own weight of attraction. We pick randomly from them with a probability that is proportional to the sum of the attraction levels. Toursits come back to the source after having visited all the attractions. The speed that people are assigned depend on the density present in a road

$$if \rho = \begin{cases} \geq \rho_c & v_{dt} = v_c \\ \leq \rho_c & v_{dt} = (1 - \frac{\rho}{\rho_c})v_{free} \end{cases} \quad (3.22)$$

where $\rho = \frac{count}{L}$ is the ratio between number of agents on a road and its length; ρ_c and v_c are the critical values for density and speed (they change if we consider agents as pedestrians or cars); v_{free} is the free speed of that road that can change according to the transport mean considered. For the simulations we used $\rho_c = 4.0m^{-1}$ and $v_c = 0.15m/s$. These parameters are calibrated in [90] and are meant to apply fundamental diagram-like congestion. Thanks to the best path algorithm, each agent knows the path to take to reach the destination. Indeed, during the initialization of the simulation, the best path between all pairs of nodes of the cartography is computed and the weight given by the following equation:

$$W_{road} = L \left(\alpha_{we} + \alpha_{speed} \frac{v_{MIN}}{v_{free}} + (1 - \alpha_{we} - \alpha_{speed}) \frac{\frac{1}{cnt} - \frac{1}{cnt_{MAX}}}{\frac{1}{cnt_{MIN}} - \frac{1}{cnt_{MAX}}} \right) \quad (3.23)$$

where L is the length of the road; v_{free} is the free speed; v_{MIN} is the minimum traveling speed; cnt , cnt_{MIN} and cnt_{MAX} are respectively the number of people passing along that road, the minimum and the maximum number of people passing along any road.

α_{speed} and α_{we} are "buttons" that allow us to consider or not each of the three terms of equation 3.23 in the calculation of weight. The meaning of the terms are:

1. the first term is based on the length of the road; then, in this case, the best path is equivalent to the shortest path in terms of meters. It is activated by the coefficient α_{we} .
2. the second term is based on a combination of length and speed. Under the same length, the path that appears shorter is that with higher velocity. This term is activated only by α_{speed} coefficient.
3. the last term is based on a combination of length and occurrences. In this case, instead, under the same path length, the shorter is the most popular.

This weights are useful for example to model toursits, where in the case of Venice is known for example that the most popular roads are the most walked. To simulate just the flows of tourist we generate people from the station of Papadopoli and the train station, plus a fake Vaporetti station as highlighted in fig. 4.29.

3.4.2 Ward Algorithm

To unravel the structure of flows we decide to use Ward algorithm. Ward algorithm, is a hierarchical clustering method whose goal is to preserve a hierarchy of signals [94]. It finds a not pre-determined number of clusters as it solves approximately an NP-problem as explained here below, it is data driven. We chose this hierarchical algorithm that does not need prior knowledge in the number of clusters, since we do not know them in advance and we do not know the importance of flows in advance and so we need a procedure to highlight both the aspects. The hierarchical structure of flows is expected since correlations driven by attractions are expected.

Let be, $O = \{o_1, \dots, o_n\}$ the set of objects to be clustered. Let $A_{iter} = \{\alpha_1, \dots, \alpha_{cluster_{iter}}\}$ be the set of clusters at iteration iter. Let $P = \{O_i : \bigcup O_i = O\}$ a partition in clusters of O, and $C = \{P\}$ the set of all possible subsets of O. Let O be embedded in a Hilbert (important for distance concepts) vector space (containing the flows vector), whose coordinates are $v_t^{o_i}$ for o_i . We call the metric in such Hilbert space d_W to be the euclidean l^2 -norm:

$$d_W(o_i, o_j) = \sqrt{\sum_t (v_t^{o_i} - v_t^{o_j})^2} \quad (3.24)$$

We define $Centers_{iter} = \{v_t^\alpha : \alpha \in A_{iter}\}$ the set centers at iteration iter. $Centers_{iter}$ is a subset of $Centers_{all} = \{v_t^\alpha : \alpha \in \bigcup A_{iter}\}$ that is the set of all possible subsets available with data at hand. The clustering algorithm is an optimization algorithm that focuses on minimizing in the set of all possible partitions the distance between centers of centroids by solving $\min_C \sum_{center_1, center'_2 \in C} d_W(v_t^{center_1}, v_t^{center'_2})$. Let represent the coordinates of each cluster center at each iteration in an abbreviated form as the tuple $(i, iter + 1) = v_t^{center_i, iter+1}$. Then $d_W((i, iter + 1) \cup (j, iter + 1), (k, iter + 1)) = d_W(v_t^{center_i, iter+1} \cup v_t^{center_j, iter+1}, v_t^{center_k, iter+1})$ represents the distance at iter+1 between the new center that is the union of i and j and center k.

We have $d_W((i, iter), (k, iter)) = d_W(v_t^{center_i, iter}, v_t^{center_k, iter})$ the distance between center i and center k at iter. The algorithm finds a solution accepting iteratively those configurations whose:

$$\begin{aligned} d_W((i, iter + 1) \cup (j, iter + 1), (k, iter + 1)) &= d_W((i, iter), (k, iter)) \\ &\quad + d_W((j, iter), (k, iter)) \\ &\quad + d_W((i, iter), (j, iter)) \\ &\quad + |d_W((i, iter), (k, iter)) - d_W((j, iter), (k, iter))| \end{aligned} \quad (3.25)$$

The typical output of this algorithm is a dendrogram that shows the hierarchical separation that occurred in the algorithm 4.4. We used it in the case of Venice 4.4. In our case O is the set of cameras mounted in the city of Venice that measure flows of people v_t^α in time span of one hour. The distance represents the difference in flows at different points in the city. Since the 1 hour time is enough to travel any trip inside Venezia walking, we expect that cameras c count similar number of people if they travel similar trips.

3.5 Another Method to Measure Monocentrism

In this section, we introduce a method to define the level of monocentrism based on mobility-derived potential fields. This approach was conceived as an alternative to the computation of the Urban Centrality Index (UCI) described in Section 3.3.3.

There, the UCI was defined in Equation 3.15 as the product of the Location Coefficient (LC, Equation 3.17) and the Proximity Index (PI, Equation 3.19). Both indices are derived from the

static population distribution observed on a given day and thus only allow a time-independent characterization of monocentricity.

In contrast, the method presented here replaces the role of population density with a potential function derived from observed mobility flows, specifically origin-destination (OD) matrices T_{ij} . This substitution allows for a dynamic view of monocentrism, enabling the measurement of centrality across different hours of the day.

While the temporal analysis of UCI is not the focus of our current study, this formulation would in principle allow for such extensions. In our case, the method is used to assess monocentricity during peak traffic hours, consistent with the problem introduced in Section 3.3.

The goal of this study is to show that the potential defined in 3.5.5 can be thought as a measure of preference defining a hierarchy that can be used in input to the computation of UCI.

Operational Procedure. We begin with a vector field defined on a spatial grid, obtained from OD matrices as described in Equation 3.33. This vector field is then decomposed using the Hodge decomposition (Equation 3.32), which separates it into gradient (potential), curl, and harmonic components. In our setting, the latter two components are negligible, allowing the interpretation of the vector field as being largely driven by a scalar potential.

Hierarchy Interpretation. The construction of a hierarchy of centers is based on interpreting the vector field components as pairwise preferences, analogously to the pairwise matrices introduced in Equation 3.28 and discussed in Section 3.5.1. Although this interpretation does not have a direct physical meaning, it provides a mathematically consistent framework, as further elaborated in Section 3.5.6.

This method was not used for the simulations in Section 3.3 because the algorithm employed to generate synthetic configurations of UCI (Section 3.3.6) was not ergodic with respect to UCI when applied to mobility-based potential fields, unlike the case with static population. This limitation arises from the complex dependency between potential and population distributions.

Conceptual Background. This method draws conceptual inspiration from a combination of insights presented in [9], [85], and [77]. Even though the pairwise preference matrix and the vector field lack a direct empirical interpretation, the potential field—derived from outgoing flows—is interpreted as a proxy for the attractiveness of each spatial cell. Intuitively, more attractive areas draw higher inflows and are thus more likely to be central.

The construction of the potential field is detailed in Section 3.5.5.

Final Remarks. The emergence of a meaningful hierarchy from this method is not trivial. The vector field aggregates global information, but each "voter" (grid cell) expresses preferences only locally, typically with respect to its nearest neighbors. Despite this sparsity, the mathematical structure remains consistent and robust.

Beyond enabling a temporally dynamic view of monocentrism, this approach offers a conceptually new way to understand urban centrality as a function of mobility needs, rather than solely static population distributions.

3.5.1 Theoretical Framework: from pairwise votes to ranking

We here define the problem and give the theoretical definitions and theorem at the base of the construction of the potential and definition of centers. As a reference for the notations we use 2.3.2.

We call V the set to be voted and Ω set of voters, $\alpha \in \Omega$ is a voter $i, j \in V$ of objects to be compared. For each voter α then we have a graph (G, V, E) that holds information about pairwise ranking in the form of a matrix:

$$T_{ij}^\alpha = \begin{cases} T_{ij} & \text{If } \alpha \in \Omega, i, j \in V \\ 0 & \text{otherwise} \end{cases} \quad (3.26)$$

This matrix is such that:

$$T_{ij}^\alpha = -T_{ji}^\alpha \quad (3.27)$$

If one considers

$$T_{ij} = \sum_{\alpha} T_{ij}^\alpha \quad (3.28)$$

In our study case $\Omega = \text{grid}$ and $i, j \in \text{cells}$, α are people that have origin in i and destination in j . The collective behavior of single preferences is encoded in the pairwise matrix T_{ij} . Properties of the matrix will define whether it is possible to obtain a rank R_i in the nodes. The nodes are the cells in the grid in this case. The ranking is a scalar function of the cells. We recover the Hodge decomposition to show that the only way to have a scalar function on the node is to have harmonic and curl components $= 0$.

3.5.2 Definitions

To tackle this problem we recover combinatorial hodge theory on graphs.

Definition 3.5.1. combinatorial gradient operator maps a potential function on the vertices $S : V \rightarrow \mathbb{R}$ to an edge flow $\text{grad } S : V \times V \rightarrow \mathbb{R}$ via

$$(\text{grad } \phi)(i, j) = \phi_j - \phi_i. \quad (3.29)$$

An edge flow that has this form will be called a gradient flow.

Definition 3.5.2.

globally consistent ranking (GCS) Pairwise ranking matrix that is written as $(\text{grad } \phi)$

Given $\tau(E) := i, j, k \in \binom{V}{n}(i, j), (j, k), (k, i) \in E$

Definition 3.5.3. combinatorial curl operator that maps edge flows to triangular flows by

$$(\text{curl } T)(i, j, k) = \begin{cases} T_{ij} + T_{jk} + T_{ki} & \text{if } \{i, j, k\} \in \tau(E), \\ 0 & \text{otherwise.} \end{cases} \quad (3.30)$$

In other words, the combinatorial curl takes pairwise rankings to triplewise rankings. The skew-symmetry of T , i.e. $T_{ij} = -T_{ji}$, guarantees that $\text{curl } X$ is a triangular flow, i.e.

$$(\text{curl } T)(\sigma(i), \sigma(j), \sigma(k)) = \epsilon_{ijk}(\text{curl } T)(i, j, k)$$

Where ϵ_{ijk} is the anti-symmetric tensor.

Let $T : V \times V \rightarrow \mathbb{R}$ be a pairwise ranking edge flow on a pairwise comparison graph $G = (V, E)$.

Definition 3.5.4. T is **consistent** on $\{i, j, k\} \in \tau(E)$ if it is curl-free on $\{i, j, k\}$, i.e.

$$(\text{curl } T)(i, j, k) = T_{ij} + T_{jk} + T_{ki} = 0.$$

Note that this implies that $\text{curl}(T)(\sigma(i), \sigma(j), \sigma(k)) = 0$ for every permutation σ .

Definition 3.5.5. T is **locally consistent** (LCS) or **triangularly consistent** if it is curl-free on every triangle in $\tau(E)$, i.e. every 3-clique of G .

Clearly, any gradient flow must be curl-free everywhere, i.e. the well-known identity in vector calculus

$$\Delta_1 = \text{curl} \circ \text{grad} = 0 \quad (3.31)$$

is also true for combinatorial curl and combinatorial gradient. When G is an incomplete graph, the condition that T is curl-free on every triangle in the graph will not be enough to guarantee that it is a gradient flow. The reason lies in that curl only takes into account the triangular inconsistency; but since there are missing edges in the pairwise comparison graph G , it is possible that non-triangular cyclic rankings of lengths greater than three can occur. Hodge decomposition theorem will tell us that all such cyclic rankings lie in a subspace of harmonic rankings, which can be characterized as the kernel of the combinatorial Laplacian.

3.5.3 Main Theorem

Helmholtz Decomposition Theorem. Let $G = (V, E)$ be an undirected, unweighted graph and K_G be its clique complex. The space of edge flows on G , i.e. $C^1(K_G, \mathbb{R})$, admits an orthogonal decomposition

$$C^1(K_G, \mathbb{R}) = \text{im}(\text{grad}) \oplus \ker(\Delta_1) \oplus \text{im}(\text{curl}^*) = GCS \oplus LCS \oplus HC. \quad (3.32)$$

This theorem states in particular that a ranking matrix can be decomposed as: Where GCS stands for globally consistent ranking, LCS locally consistent and HC harmonic component.

Therefore if one checks that:

- Harmonic Component is 0
- Rotor is 0

Then one obtains a consistent ranking.

Essentially what these theorems are telling, is that different regions are voting one another by sending people among them. They are telling whether they prefer some place over the other. What we need to avoid are loops both local (rotor components) and global (harmonic components) to be consistent; an example of consistency is shown in fig. [3.12](#).

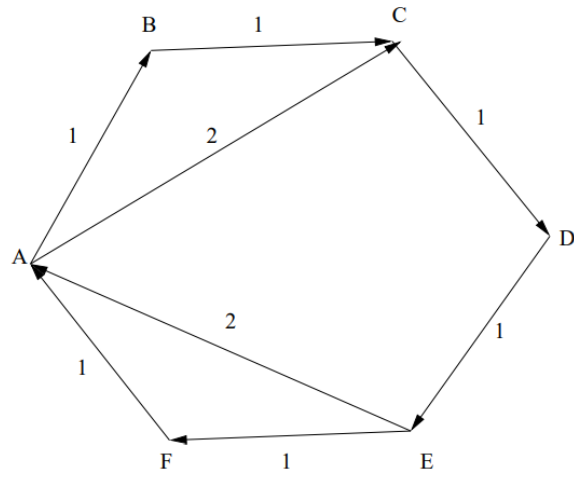


Figure 3.12: Image from [9]. This is an example of harmonic component that is inconsistent globally but consistent locally on every triangle. The inconsistency is found along the loop $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow A$.

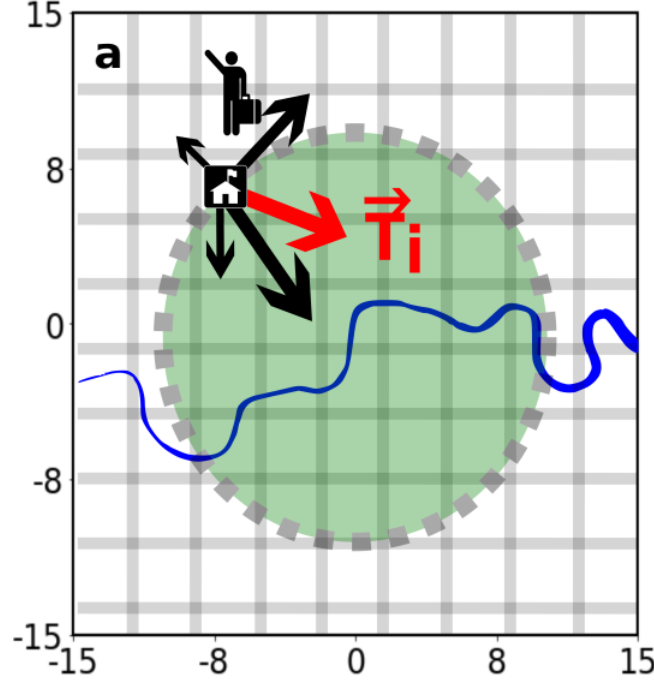


Figure 3.13: Sketch of the construction of the vector field taken from [85], where we can see from the origin cell different destinations that are vectorially added together to build the vector field.

3.5.4 Vector Field

An interesting approach to study mobility is a field theory approach and it was introduced in [85] with later expansion in [76]. In that work the authors measure flows among different grids used to tessellate the region of interest (T_{ij}). They aggregate the flows into a vector field and model it according 3.33. The vector field is obtained by summing over the possible destinations and dividing by m_i (population of the grid). If $\vec{u}_{ij}$ is the unit vector pointing from j to i , the vector field (fig. 3.13) can be written as

$$\vec{W}_i = \sum_j \frac{T_{ij}}{m_i} \vec{u}_{ij} \quad (3.33)$$

3.5.5 Potential

The potential is calculated by numerically solving the equations $-\nabla V_i = \vec{W}_i$ taking into account that The condition $\nabla \times \vec{W} = 0$ is used. For the computation of the empirical potential, we used the condition $V = 0$ in all the boundary regions of the grid and then used the forward centered discretization formula for the gradient operator starting from the city bounding box corner. In a cell i with indices (α, β) , this operation becomes:

$$\frac{dV_i}{dx} = \frac{V_{\alpha+1,\beta} - V_{\alpha,\beta}}{\Delta x} = W_{(x),\alpha,\beta} \quad (3.34)$$

$$\frac{dV_i}{dy} = \frac{V_{\alpha,\beta+1} - V_{\alpha,\beta}}{\Delta y} = W_{(y),\alpha,\beta}. \quad (3.35)$$

The procedure is iterated until all cells have been assigned a potential. We then average the resulting potentials after starting from every corner of the bounding box to decrease the noise.

3.5.6 Vector Field Potential and Pairwise Hierarchies

In this section, we construct a formal correspondence between the potential-based mobility vector field introduced by [85] and the hierarchical structure inferred from pairwise comparison matrices, as in [9].

The key property underpinning this connection is the *antisymmetry* of the pairwise flow matrix. Specifically, we aim to relate the antisymmetric structure described in equation 3.27 to the spatial components of the vector field described in equations 3.35 and 3.34.

To illustrate this connection more explicitly, we introduce a refined indexing scheme. Let i and j be monodimensional indices labeling cells on a two-dimensional spatial grid. Each grid cell indexed by i corresponds to a unique spatial coordinate (x_i, y_i) .

Consider the potential function $V : \mathbb{R}^2 \rightarrow \mathbb{R}$, from which a vector field is derived via $\vec{F} = -\nabla V$. For concreteness, we focus on the discrete derivative in the y -direction:

$$\frac{dV_i}{dy} \approx \frac{V_{(x_0, y_1)} - V_{(x_0, y_0)}}{y_1 - y_0} = T_{(x_0, y_0), (x_0, y_1)} = T_{ij}, \quad (3.36)$$

where $i = (x_0, y_0)$ and $j = (x_0, y_1)$ is a neighboring cell, i.e., $y_1 \in \partial y_0$.

This construction shows that the discrete gradient naturally defines a pairwise quantity T_{ij} that satisfies

$$T_{ji} = -T_{ij}, \quad (3.37)$$

since swapping i and j simply reverses the direction of the difference. Thus, the antisymmetry of T arises directly from the spatial embedding of the grid and the use of finite differences to approximate derivatives.

This observation provides a geometric interpretation of pairwise comparison matrices as discrete gradients of scalar fields. In this way, the tools of Hodge theory—such as decomposition into gradient (potential), curl (circulatory), and harmonic (global inconsistency) components—can be applied to analyze the mobility field structure derived from observed flows.

Chapter 4

Results

This chapter is devised as a container for all the results obtained in the thesis for all the different projects I have worked on. Each section in this chapter has at least a corresponding one in chapter 2 and chapter 3.

4.1 Mobility changes during COVID-19

In this section, we present two case studies that quantify behavioral changes in mobility due to lockdown restrictions, based on GPS data collected from individuals living in California (Section 4.1.2) and Massachusetts (Section 4.1.2), using the methods introduced in Section 3.1. The datasets used are those provided by Cuebiq and described in Section 2.1.1.

The analytical pipeline is the same for both case studies; however, for Massachusetts, we also include a stability analysis to validate the methodology. This analysis is presented in Appendix A.1.

The goal of this study is to describe how COVID-19 lockdowns perturbed mobility metrics from their equilibrium state.

The main contribution of this study is to provide quantitative insights into mobility changes that can inform improved modeling of epidemic spread. Epidemic models [6, 8] often incorporate mobility patterns parametrically at the metapopulation level, where they define link weights in the spatial network connecting communities. These models typically rely on assumptions about baseline mobility and introduce interventions as changes in contact probabilities. We suggest that empirical information on travel behavior can support the development of more realistic models in the direction of the construction of realistic synthetic trajectories. In particular, we highlight behavioral constraints relevant for generating synthetic trajectories that can be used to estimate inter-community mixing.

4.1.1 California

The first measurements we present are the radius of gyration and the geodesic distance distribution (Figs. 4.1, 4.8). From an epidemiological perspective, these metrics are particularly relevant, as reductions in individuals' spatial range are expected to decrease the rate at which an epidemic spreads across—potentially distant—locations. All the results that we show in the next two paragraphs are executed with $t^{burst} = 1hour$.

The results show that, during lockdown periods, most trips were concentrated near the trajectory's center of mass (within 2 km), compared to the pre- and post-lockdown phases. Even outside lockdowns, the most frequent trips remained short (under 1 km), as seen in the distance distributions (Figs. 4.1, 4.8).

This pattern, together with the returner–explorer distribution (Fig. 4.2), suggests that people became significantly more likely to make short trips and return to familiar locations—behaving more as “returners” than “explorers.” This tendency is further confirmed by the visitation frequency and discovery rate analyses (Fig. 4.5), where lockdown periods are marked by a reduction in the rate of discovering new places and an increase in routine visits to already-known locations.

Although mobility patterns were spatially constrained during lockdown, their geometric distribution retained an approximately invariant two-dimensional shape (Fig. 4.3), with activity peaks shifting closer together. This indicates that individuals generally stayed nearer to their homes, minimizing the spatial extent of their movements. However, the overall structure of mobility remained consistent, supporting the interpretation that this structure reflects essential activities—such as commuting or grocery shopping—that continued under restrictions.

This simplification of trajectories is also evident from an information-theoretic perspective (Fig. 4.7). Across all entropy measures considered, the information content of mobility trajectories is consistently lower during lockdowns. This may result from increased regularity in movement—i.e., more periodic sequences of visited locations—or from a decrease in the number of distinct places visited. The latter seems more likely, as indicated by the shift in entropy distributions normalized by the number of visited locations. This shift suggests that the Lempel-Ziv compression algorithm (Fig. 4.7d) extracts less information from trajectories, pointing to reduced uncertainty in the sequence of visited locations.

Finally, we examine differences in mobility across population groups, categorized by rurality (defined as areas with population density below 500 people/km²) and by the political affiliation of the local government. While rurality appears to have limited impact on the weekly radius of gyration, political affiliation does: individuals living in Republican-governed counties exhibited higher mobility during lockdown periods (Fig. 4.6).

Together, these findings provide a comprehensive snapshot of how mobility patterns in California and Massachusetts changed under non-pharmaceutical interventions (NPIs), highlighting shifts in the size, shape, and uncertainty of movement, as well as their socio-political correlates.

We observe that the patterns are broadly similar in both Massachusetts and California, suggesting comparable effects of the lockdowns across these two states.

California	N_c	T_c	$\langle T_c \rangle_p$
15 January - 15 March	1,009,464	164,994,721	163
15 March - 15 May	490,673	12,701,431	26
15 May - September	460,654	32,096,311	70

Table 4.1: Main information about the California preprocessed dataset with columns N_c number of users, number of total points T_c and average points per each trajectory $\langle T_c \rangle_p$

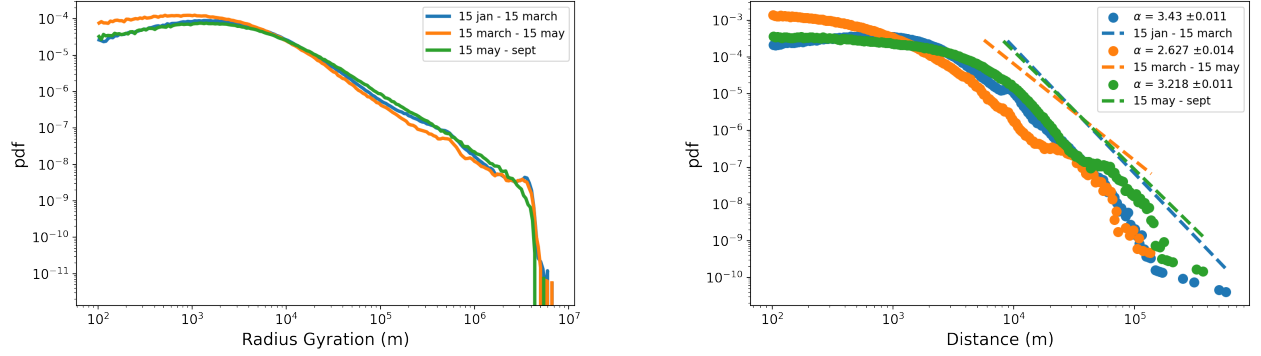


Figure 4.1: California: On the left: Distribution of radius of gyration among in three different periods. The lockdown curve is above the other two curves for distances below 3 km and below up to 1000 km, meaning that medium-distance mobility is redistributed on the short-distance one, while preserving long trips mobility. On the right: Distribution of distances among successive stop points in three different periods and their corresponding power laws. Prevalence of short distance travels and long tail can be observed. Time threshold = 1 h

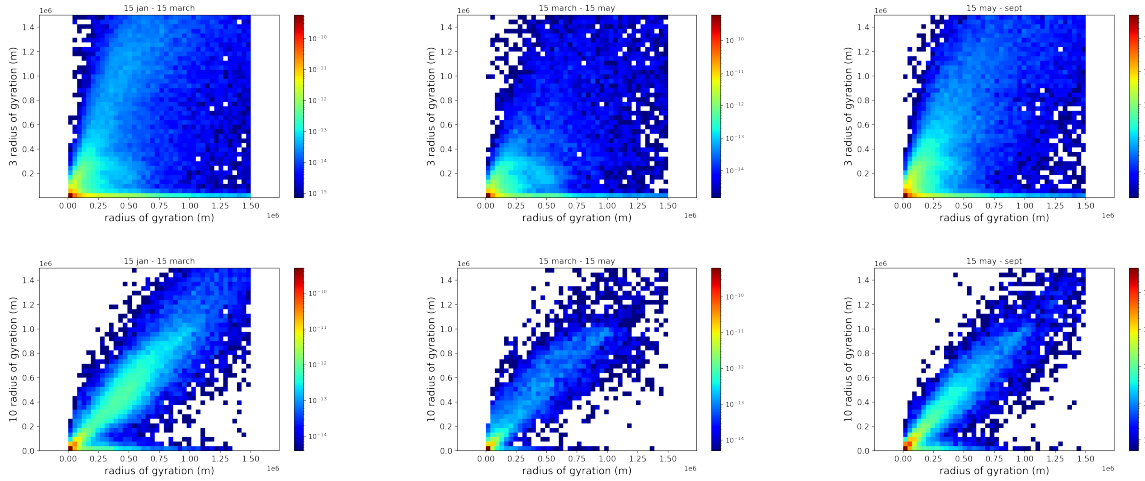


Figure 4.2: California: Distribution of (radius of gyration, k-radius of gyration) in three different periods. From top to bottom $k = 3, 10$. From left to right period of measure (15 January - 15 March, 15 March - 15 May, 15 May - September). Time threshold = 1 hours.

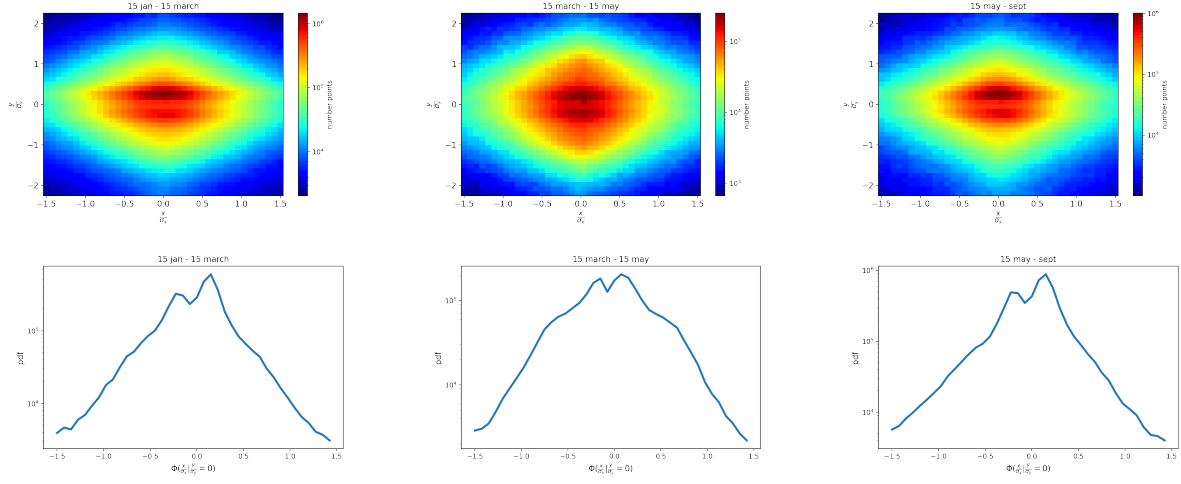


Figure 4.3: California: Distribution of positions of people stops' points in their intrinsic reference frames re-scaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). The bottom figures are the conditional probability fixing the coordinate $y = 0$. Case of time threshold of 1 hour

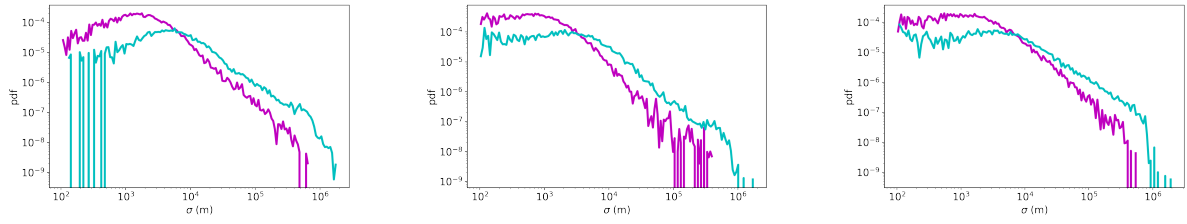


Figure 4.4: California: Distribution of variances of trajectories stops' points in their intrinsic reference frames re-scaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). Time threshold of 1 hour.

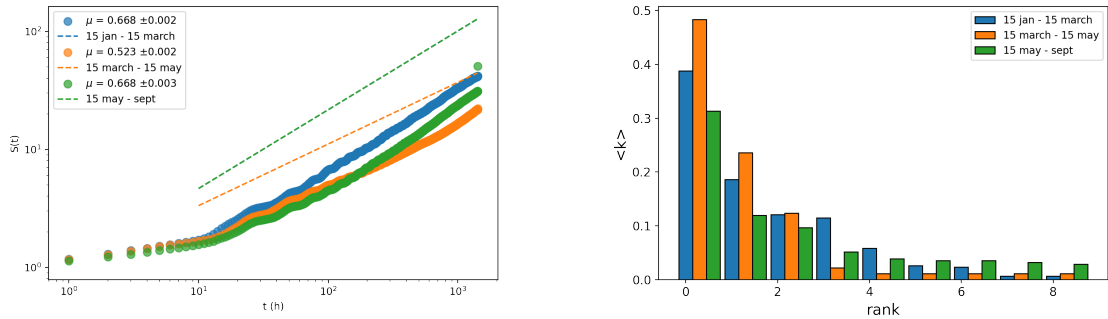


Figure 4.5: California: On the left: Average number of stop points visited in time. During lockdown, the rate of discovery is lower than the other two periods. On the right: Distribution of frequency of visits of first ten preferred places for each user. During lockdown, the first two ranked visited places occupy 70 % of the visitation pattern, against 50% and 40% of the before and after periods. Case of time threshold of 1 hour

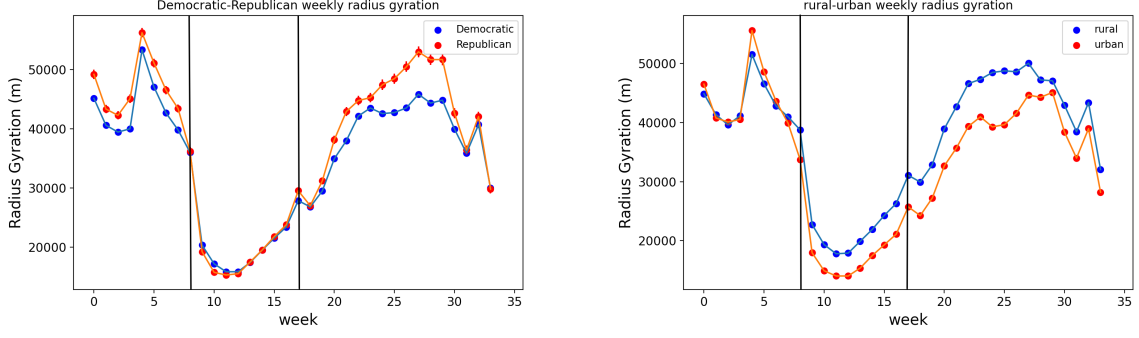
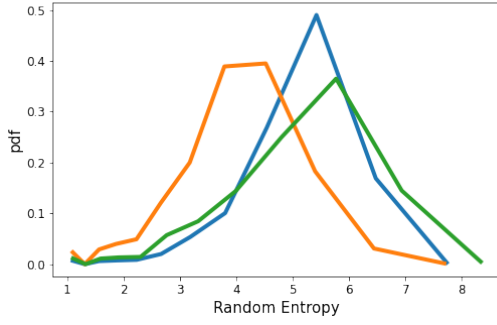
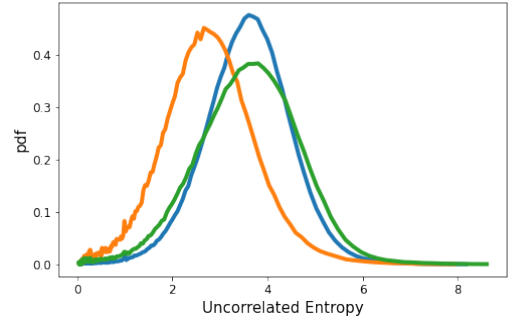


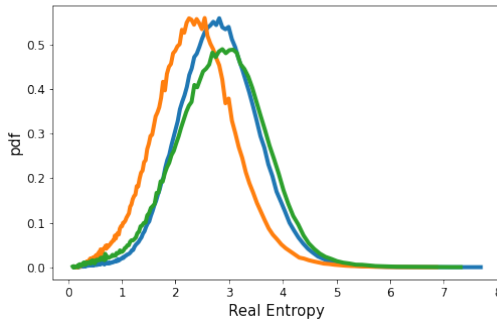
Figure 4.6: California: On the left: Average radius of gyration per week for people characterized by the party governing their home county. On the right: Average radius of gyration per week for people characterized by the level of rurality in their home. The different period, (pre,during,post)-lockdown, are separated by vertical lines. The party governing is not very significant in the prediction of the radius of gyration. On the other hand, the rurality level shows that people living in rural areas have higher mobility both during and after lockdown with respect to people living in urban areas. Case of time threshold of 1 hour



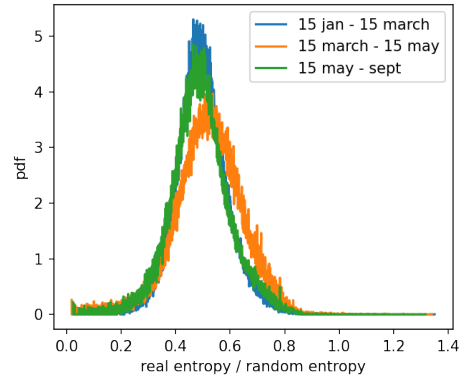
(a) Random entropy. During lockdown, the distribution shifts left, indicating reduced trajectory uncertainty.



(b) Uncorrelated entropy. Also shows a leftward shift during lockdown, consistent with simpler mobility patterns.



(c) Real entropy. Confirms that lockdown trajectories are less uncertain on average.



(d) Entropy ratio: Random/Real. During lockdown, real entropy approaches random entropy.

Figure 4.7: Entropy-based uncertainty analysis of mobility trajectories in California across three periods. All entropy measures show a significant reduction in trajectory uncertainty during the lockdown. We suggest the reduction is mainly driven by the number of visited places and not by the uncertainty of the pattern

4.1.2 Massachusetts

In this section we report the results for the case of Massachusetts. The considerations about the results are very similar to the case of .

Massachusetts	N_c	T_c	$\langle T_c \rangle_p$
15 January - 15 March	181,788	28,782,953	158
15 March - 15 May	140,913	12,701,431	90
15 May - September	141,285	32,096,311	227

Table 4.2: Main information about the Massachusetts preprocessed dataset with columns N_c number of users, number of total points T_c and average points per each trajectory $\langle T_c \rangle_p$

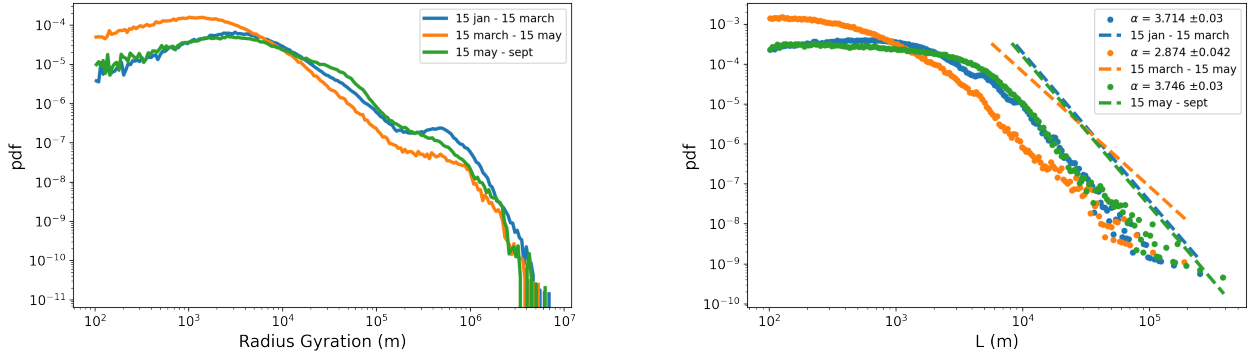


Figure 4.8: Massachusetts: On the left: Distribution of radius of gyration among in three different periods. The lockdown curve is above the other two curves for distances below 5 km, and below up to 1000 km, meaning that medium-distance mobility is redistributed on the short-distance one, while preserving long trips mobility. On the right: Distribution of distances among successive stop' points in three different periods and their corresponding power laws prevalence of short distance travels and long tail can be observed. Time threshold = 1 h

Table 4.3

	15 January - 15 March	15 March - 15 May	15 May - September
15 January - 15 March	1	0.8	0.8
15 March - 15 May		1	0.5
15 May - September			1

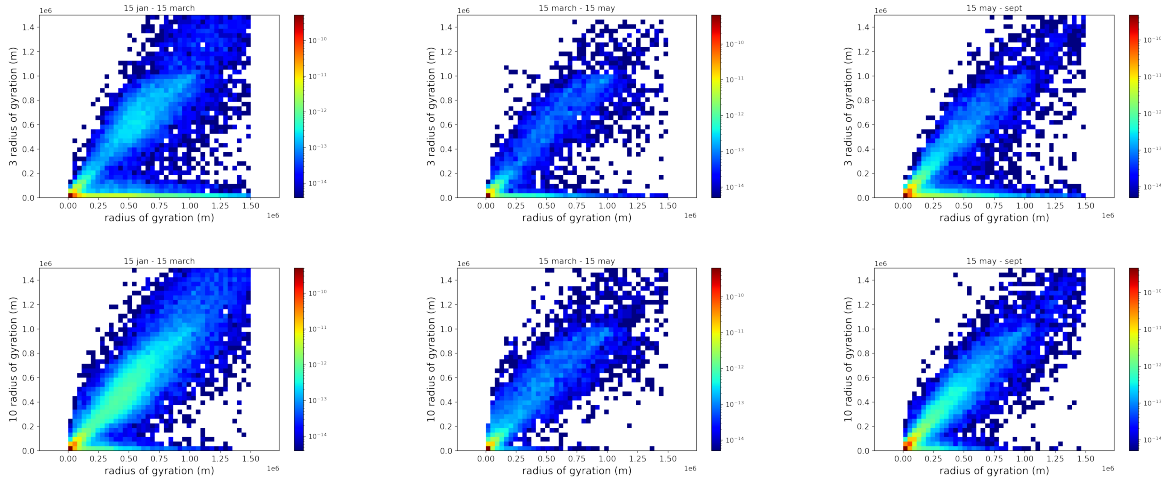


Figure 4.9: Massachussets: Distribution of (radius of gyration, k-radius of gyration) in three different periods. From top to bottom $k = 3, 6, 10$. From left to right period of measure (15 January - 15 March, 15 March - 15 May, 15 May - September). Time threshold = 1 hours

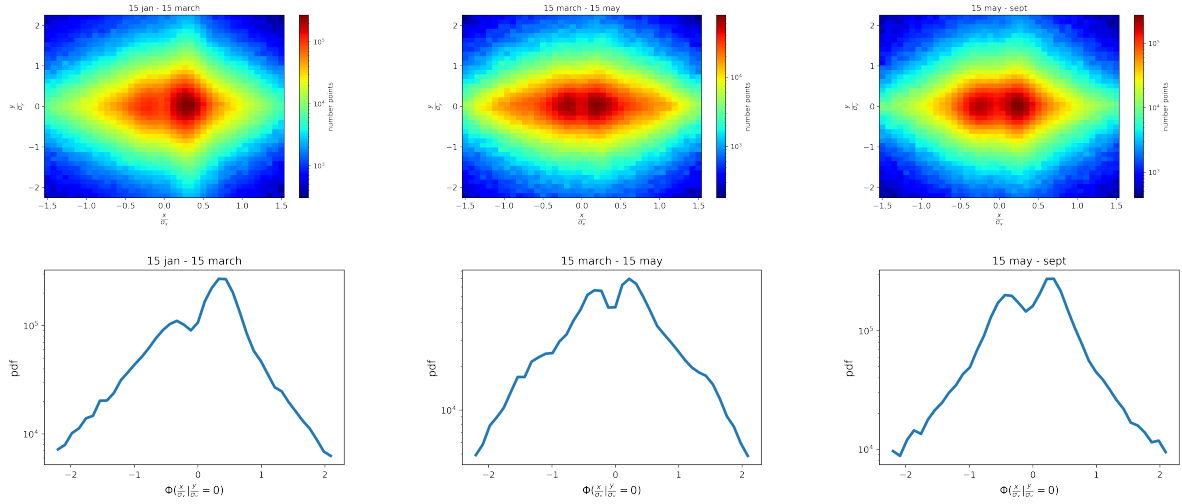


Figure 4.10: Distribution of positions of people stops' points in their intrinsic reference frames rescaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). Bottom figures are the conditional probability fixing the coordinate $y = 0$. Case of time threshold of 1 hour

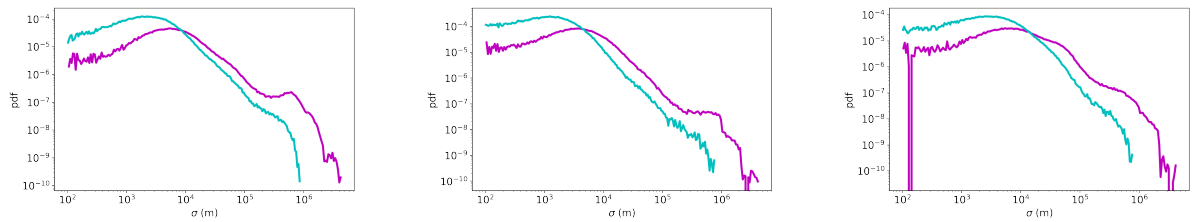


Figure 4.11: Distribution of variances of trajectories stops' points in their intrinsic reference frames rescaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). Time threshold of 1 hour.

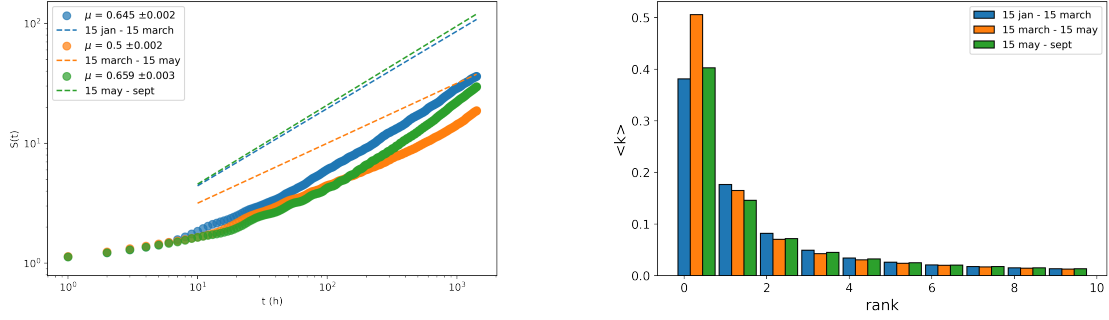
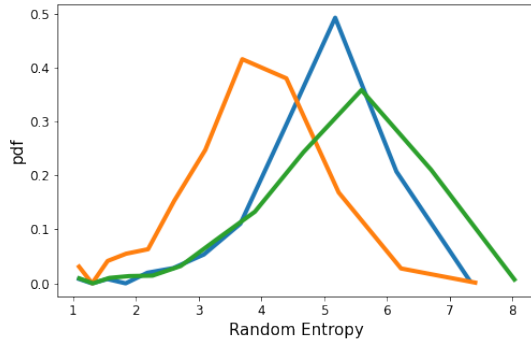
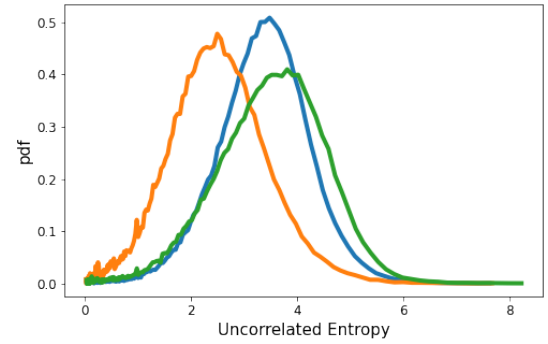


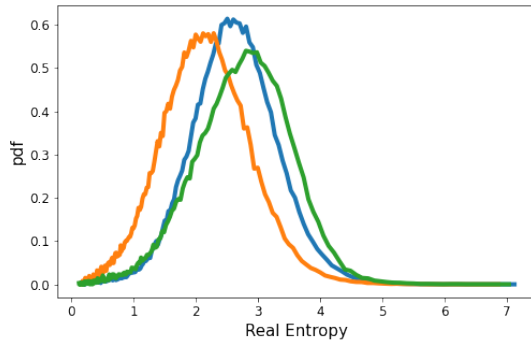
Figure 4.12: On the left: Average number of stop points visited in time. During lockdown, the rate of discovery is lower than the other two periods. On the right: Distribution of frequency of visits of first 10 preferred places for each user. During lockdown the first two ranked visited places occupy 66 % of the visitation pattern, against 50 % of the before and after periods. Case of time threshold of 1 hour



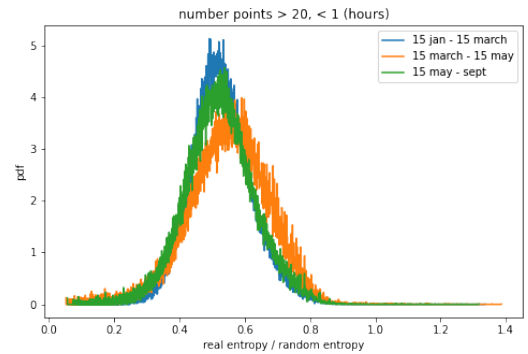
(a) Random entropy: lower during lockdown, indicating reduced uncertainty.



(b) Uncorrelated entropy: also shows a leftward shift during lockdown.



(c) Real entropy: confirms the trend of less uncertain trajectories during restrictions.



(d) Entropy ratio (Random/Real): closer to 1 during lockdown, suggesting fewer visited places drive the reduction in uncertainty.

Figure 4.13: Massachusetts: Entropic measures of trajectory uncertainty across three periods. All entropy measures decrease during lockdown, reflecting more predictable and spatially constrained mobility behaviors. These measure are very similar to the ones obtained in 4.7, highlighting homogeneous response in different populations.

4.2 The multimodal structure of the road network in Bologna

In this section we present the results of the analysis introduced in sect 1.3 whose data are described in 2.2.1 and whose goal and methodologies are listed in 3.2. We stress that the aim of this study is to define an aggregated metric for congestion in the Bologna road network, based on trajectory data collected via GPS from mobile phones. This objective is motivated by a smart-city perspective, where it is crucial to quantify the congestion state of the road network using available data sources. The contribution of this study relies on the identification of sub-networks for which we can extract a congestion index. This congestion index in particular allows us to identify different types of mobility, in different parts of the network, and understand how their mixing produces congestion patterns.

We first describe the paths properties associated to the GPS trajectories in , we continue by describing the multimodality structure of the road network and the relative traffic index associated to each layer in 4.2.2

4.2.1 Statistical properties paths

In the study of the statistical properties of urban mobility assuming a stationary state for the city, it is useful to consider the path length and the travel time distributions. In the case of MDT paths one has the possible bias that the paths could not consider the whole OD journey but the extensive use of smartphones maintains the main features of the distributions. In Fig. 4.14 we consider the MDT path length distribution for the all days in the dataset in the considered area. We consider just paths that have a duration below 1.5 hours whose length is less than 10 km to avoid comparing taxis, buses and other transportation means that travel all the day in the city of Bologna, since our goal is to study the citizen mobility. The path length distribution show an exponential tail for trips $l \leq 4$ km ($P(l) \propto e^{-\lambda l}$, $\lambda \sim 0.31 km^{-1}$) with a cutoff for very short trips where the contribution of private car is less relevant. The trip length distribution could be affected by the use of mobile phones, but the data suggest that the duration of the mobile phone activities cover long time interval (see time duration distribution Fig. 4.14). The exponential distribution is consistent with a MEP for the mobility in the city [19], but, interestingly, we have a peak in the path length distribution at $\simeq 6$ km that could be related to the mobility demand along the highway which connects the city along the east-west direction. Indeed the paths on the highway has a length distribution that substantially differs from an exponential (see supplementary material). The collapse of the distributions for the different days denotes the periodic character of the urban mobility and of the mobility demand. The initial exponential gives a characteristic path length $\simeq 3.2$ km that could be interpreted as an average mobility demand in the city of Bologna not related to a particular transportation mean. As we shall show in the next section we have a hierarchical structure in the Bologna road network that emerges using path features as the average velocity, but the exponential distribution in the path length suggests that the mobility demand is almost independent from the underlying hierarchical structure of the road network. However if we are able to follow the device for a sufficiently long time, this structure can affect the path duration distribution since a single path may contain different types of mobility. Therefore the detection of the underlying multiscale structure of the road network is a key issue to explain the features of the mobility paths and to highlight the heterogeneous usage of the road network[103]. We have considered the duration time distribution of the MDT paths. Travel time has been proposed as 'energy' cost of the mobility [72, 39] and an exponential distribution is expected from the application of a Maximum Entropy Principle (MEP)[19]. However the MDT paths suggest that a power law is suitable to model the distribution (see bottom Fig. 4.14). The power law distribution is the

same for all the days in the dataset. A possible explanation can be related to the possibility offered by the MDT dataset to follow a single mobile device for a relative long time (up to 1.5 h), so that the path duration has the possibility to record the complexity of individual mobility. Under this point of view the path duration may also reflect the strategy in the use of the different road sub-networks[103], the use of time during short stops or even the effect of a multimodal mobility (e.g. car and pedestrian). We remark that the exponential behavior of the path length distribution is consistent with a MEP with the constraint that the average length exists: i.e. the urban mobility can be simulated by random walk and the only relevant information is the average path length that has the role of 'mobility energy'. The transportation means have to be convenient to travel this distance.

To verify these assumptions, we studied the variability of the velocity along a single path so that different pieces of the trajectory could be classified in a different way: i.e. they could be attributed to different sub-networks. In Fig. 4.15 we show the result of assigning each point of a path $x_\alpha(t) \forall t \in T_\alpha$ to the class whose average velocity is closest one to the actual velocity. We observe that the paths in the class 0 (slowest class) and class 3 (fastest), are not mixed, whilst the paths in the classes 1 and 2 contain many pieces that could be attributed to the other classes. This is consistent with the persistence of the power law distribution for the path duration of the 1 and 2 class, whilst this is not the case for the paths in the classes 0 and 3 that show an exponential distribution consistently with the MEP (see Fig. 4.14 bottom right). It is noteworthy that this mixing could be not a switch of the paths to another sub-network, but it is actually as if they changed and we suggest that these transitions are partly due to traffic conditions that perturb normal jump frequencies between different layers ([103]). The velocity variability that affect the path duration distribution of the classes 1,2 may induce a deviation from exponential behavior of the duration distribution. We expect the velocity variability is mainly the consequence of the traffic effects at macroscopic level. The MDT paths in slowest class 0 have a small velocity variability probably due to physical limit of the transportation means in the corresponding subnetwork, and the paths in fastest class 3 are mainly related to the highway traffic whose average velocity changes according to traffic conditions (see Fig. 3.4), but the velocity along single path does not vary. Daily representation of these results are present in the appendix A.2.

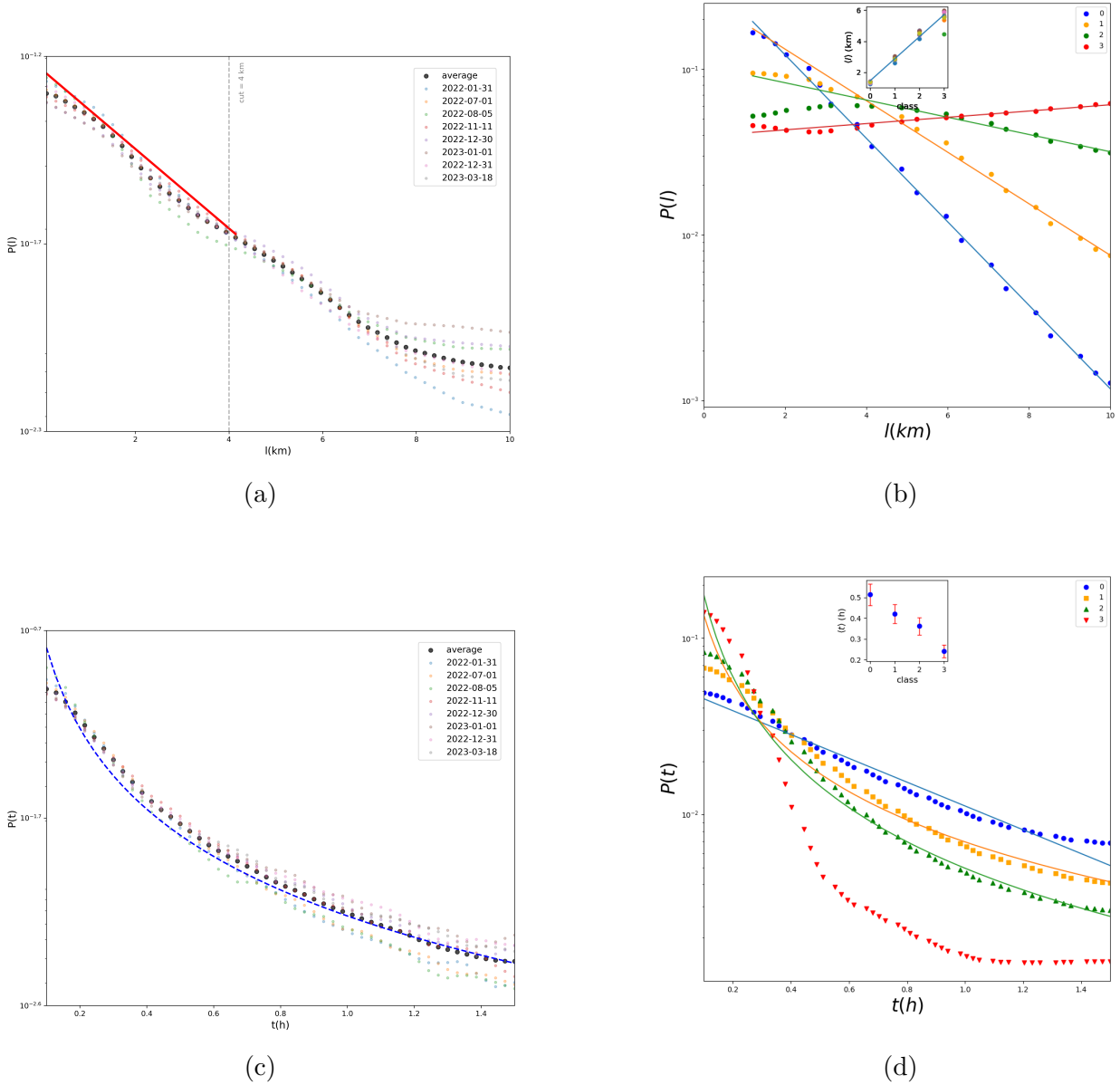


Figure 4.14: Distribution of MDT path lengths (top plots) and of the path duration distributions (bottom plots): we introduced a cut off of 10 km for the path length and 1.5 hour for the duration to mainly consider urban mobility. 4.14a: global path length distribution. An exponential fit $P_{average_days}(l) \propto e^{-l/\bar{l}}$ is proposed up to a path length of 4 km (continuous line) with a characteristic length $\bar{l} = 3.8$ km for all the days. The distribution behavior for longer paths suggests the presence of heterogeneity in the urban mobility. The corresponding path length distribution for the mobility classes (4.14b). We plot the path length distributions for the classes 0,1,2 together with an exponential fit (continuous lines) for the distribution tails and the corresponding characteristic lengths are reported in the inset. In the cases of class 1 and 2 we fit respectively $l \geq 3km$ and $l \geq 4km$. The path duration distribution for the daily MDT paths is shown in 4.14c and a power law fit $P_{average_days}(t) \propto t^\alpha$ is proposed (continuous line) with exponent $\alpha = -1.36$ (see also Fig. A.7 in the Appendix A). In 4.14d we show the path duration distributions after having disaggregated the paths into four different classes (see section 3.2). We observe that an interpolation by a power law distribution $P_{average_days}^{class}(t) \propto t^{\alpha_{class}}$ is suitable for the classes 1,2 (continuous line) with exponents $\alpha_1 = -1.67$, $\alpha_2 = -2.01$. The slowest class 0 and the fastest class 3 distribution are well interpolated by an exponential distribution ($P_{average_days}^0(t) \propto e^{-t/\bar{t}}$) with characteristic time $\bar{t} \sim 0.55h$ for the class 0 and with characteristic time $\bar{t} \simeq 10min$ for the class 3.

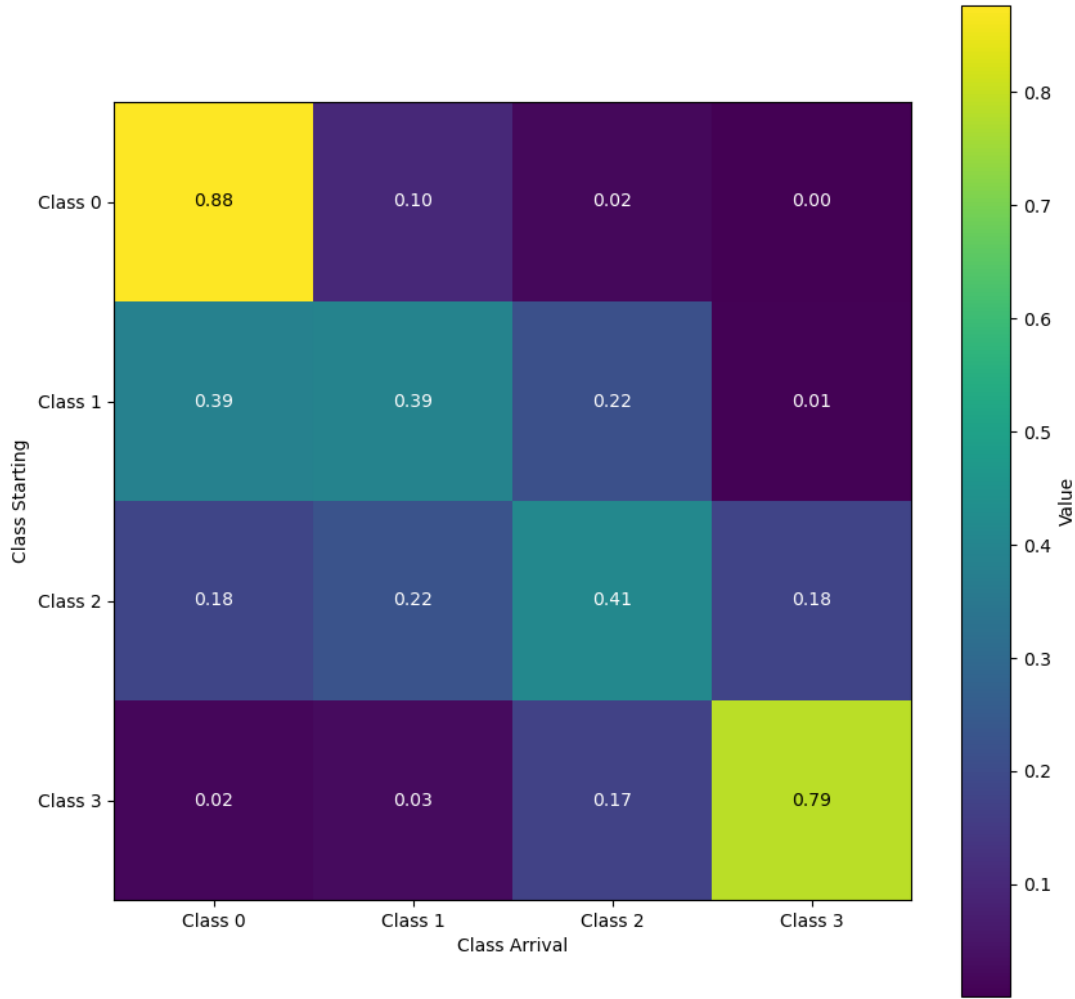


Figure 4.15: Distribution of the single path pointwise velocity among the different classes. We observe that paths of the classes 1 and 2 contain points classified in different sub-network showing their heterogeneity. This fact can explain why we observe power law distributions plotted in Fig. 4.14.

4.2.2 Speed Distribution and Multi-Modality Network

The idea behind homogeneous sub-networks, is that roads can be grouped by their functionality in hosting different speed, that is different modalities. This functionality emerges from the clustering algorithm and should be distributed accordingly a Gaussian law, hinting that the function of the sub-network is one, represented by the speed, and is noised by actual realizations of traffic that are completely random. That is not entirely true as we can see that we are not able to separate pedestrian mobility with cyclistic mobility as it can be seen from the double modality in the slowest distribution (fig. 3.4).

We consider now these distribution as benchmark of free flow distribution and look at the effect on these distribution by the re-classification induced by the hierarchical assignment. As a first measure for traffic we show the dependence of speed to the number of people in the road network fig. 4.16.

We average speed drops in 4.16 the number of people, therefore the fundamental diagram is expected [75] to be observed. To see why, we look at the perturbations to the distributions on the speed for each class. In fig. 4.17 we show on average how of the distribution in speed transforms after the hierarchical algorithm is applied. This transition of speed contributes to the drop of speed in 4.16 and generates drops of speed during peak traffic time in the middle classes sub-network where the mixing of classes in the trajectory is abundant as can be seen in fig. 4.15. As we can see from the average over all days of the curve of the traffic index showed in 4.18, we can see that for the fastest roads there are two peaks (with the third one midnight). To compare different traffic levels among them, we apply a gaussian filter to $\Gamma_k(t)$ in 1 d to smooth the signal and choose the highest two peaks and compute the average for them. This measure describes the severity of traffic in the given day for each class and allows to quantify the average value for each sub-network 4.19. Daily representation of these results are present in the appendix A.2.

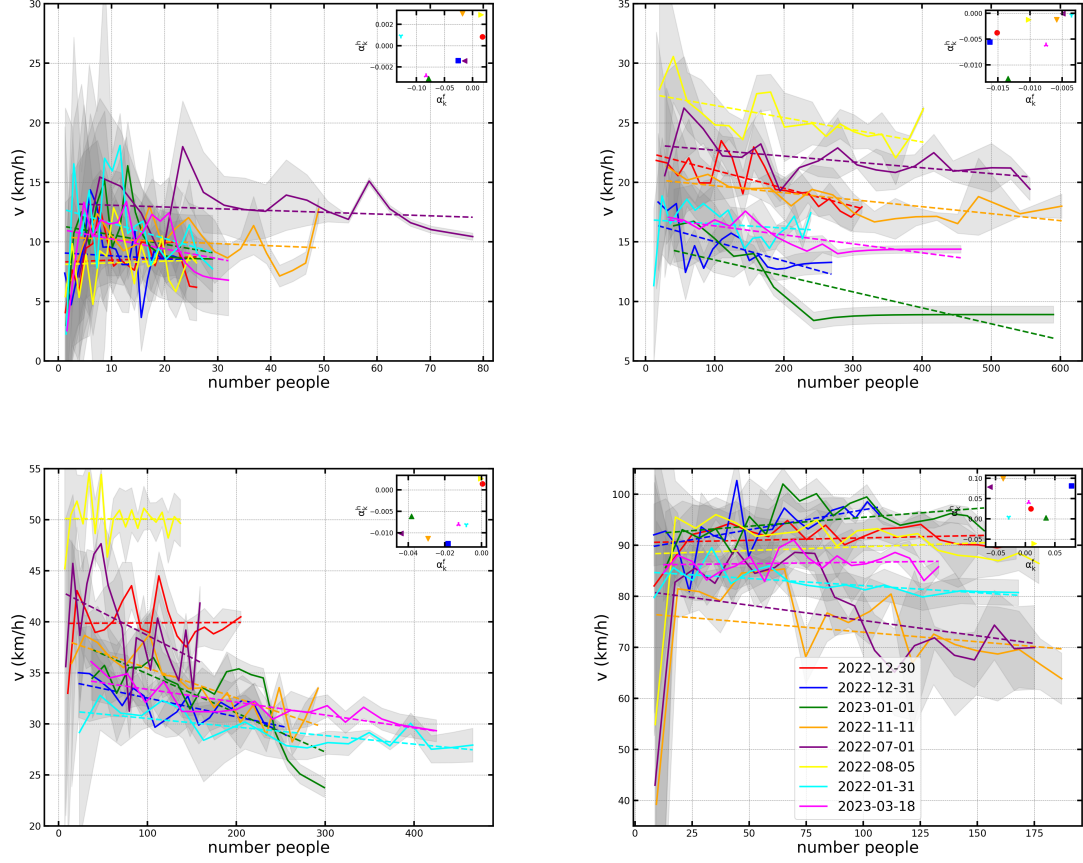


Figure 4.16: From Top left to bottom right the speed of the sub-network in terms of the number of people for the re-classified trajectories. We see that the reclassification generates trajectories from the slowest class to the second slowest and the second slowest and the second highest exchanging leaving almost unchanged the quickest (Right) Average velocity per quickest class for all different days in the dataset as a function of the number of people in the road sub-network. There is no drop of speed. In the inset the linear coefficients of the fit. They show that in the class 0, no traffic is measured, while in the other classes there is. Of particular interest in the bottom right we see that 30,31-12-2022 and 01-01-2023 do not show traffic patterns.

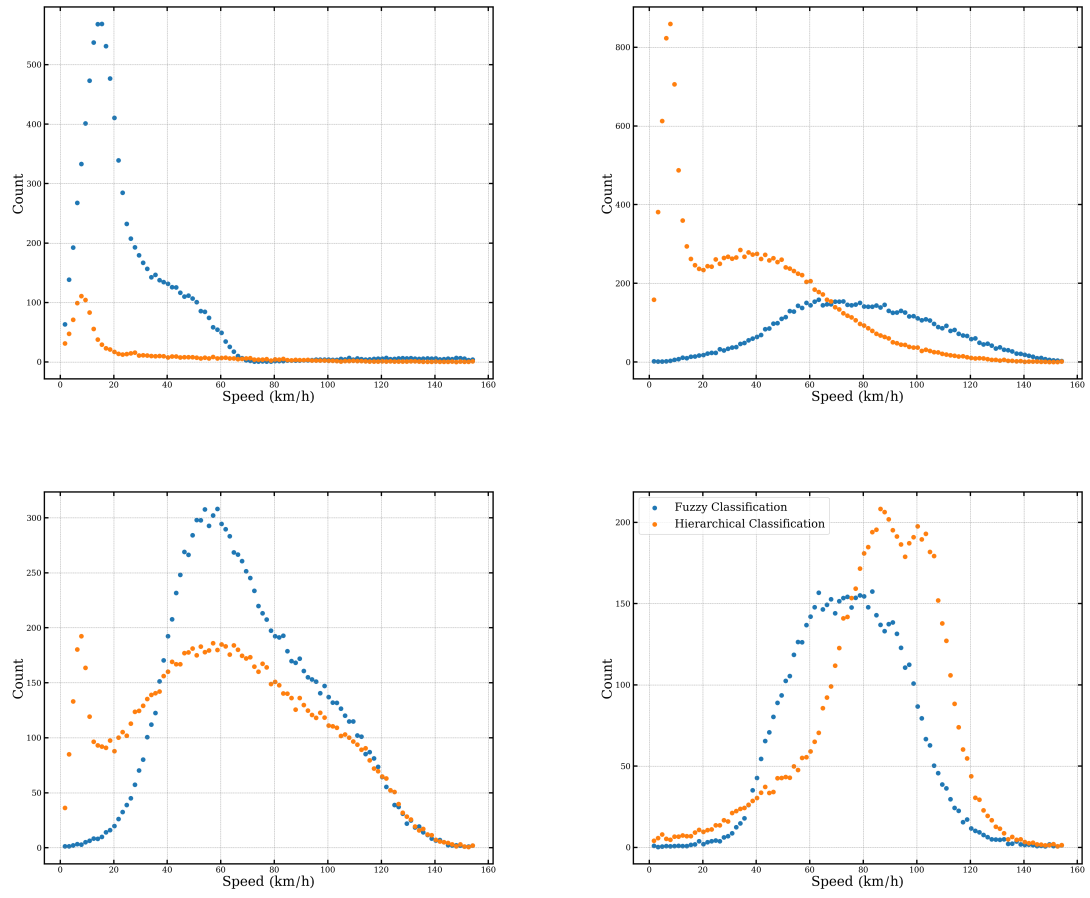


Figure 4.17: From top left (slowest class) to bottom right (quickest) we show the average change of the distribution in speed given by the hierarchical reassignment of the trajectories. We see the massive reclassification of the slowest class into the second slowest and

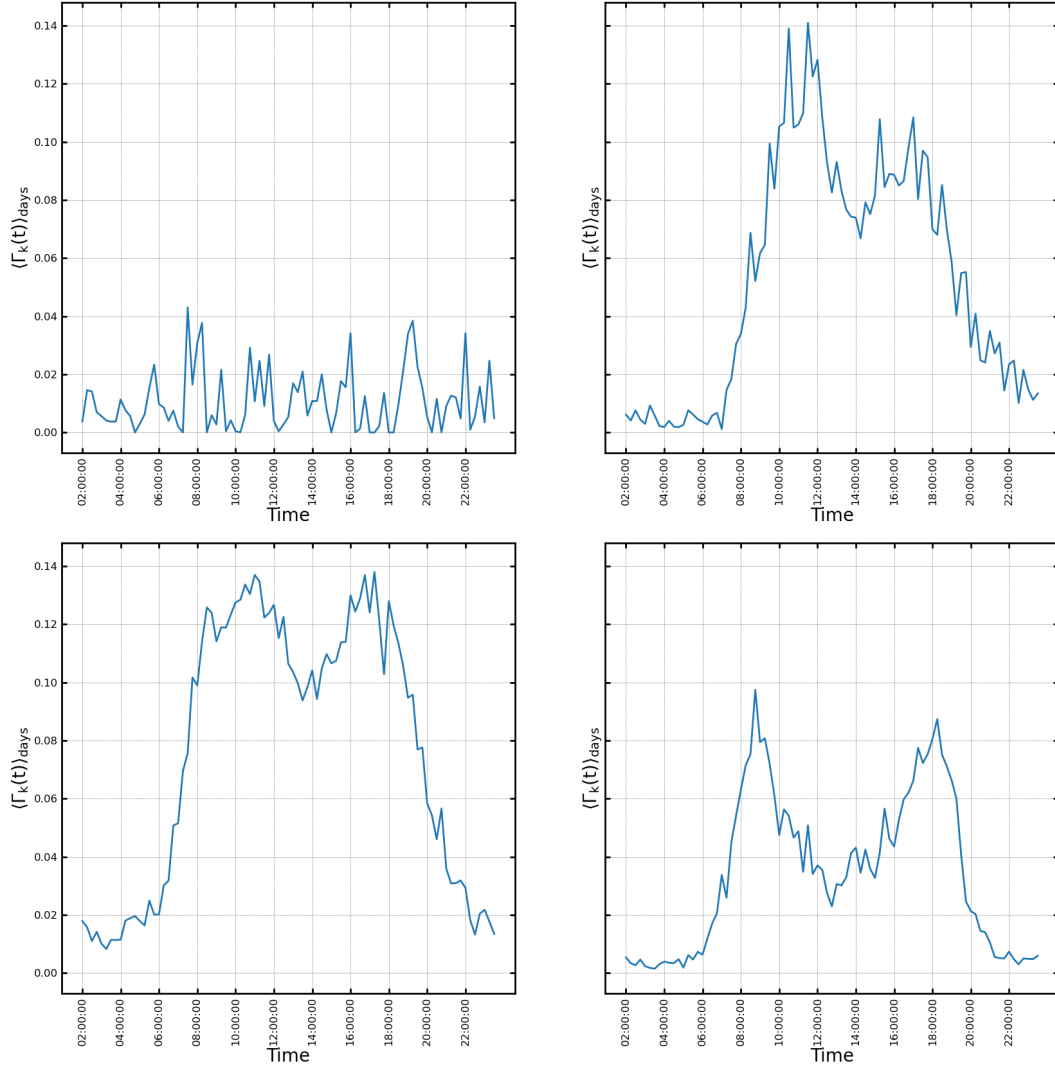


Figure 4.18: From top left to bottom right is represented the traffic index $\Gamma(t)$ averaged over all days for the roads in the hierarchical sub-network in increasing order of average speed. We can see a neat signal in the rush hours in the quickest network at 8:00 and 18:00. Peaks of traffic are measured shifted toward later times also in the second slowest and second quickest class, while noise is present in the first class that have lower values of the index.

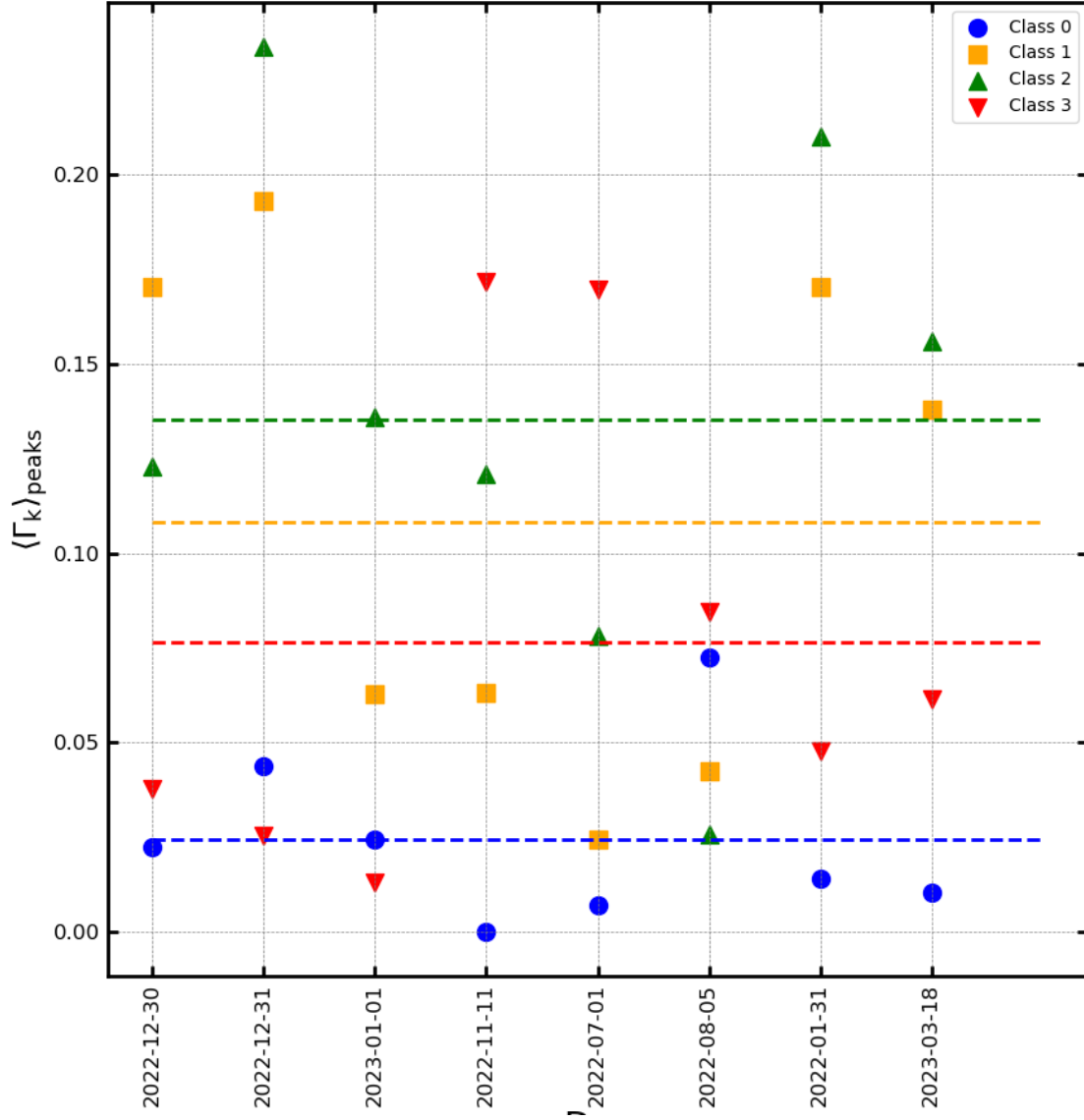


Figure 4.19: Average of traffic index ($\langle \Gamma_k \rangle_{peaks}$) computed with the two highest peaks obtained after a gaussian filter with $\sigma = 3$. Since the index is inversely proportional to the speed of the layer, the signal traffic is bigger for the second highest, and kept for second slowest. This suggests that the relative drop of flows due to traffic regimes, are captured in the two highest speed network.

4.2.3 Sub-Networks

We show here the sub-networks obtained thanks to the algorithm after having applied the hierarchy on them 4.20. This reconstruction shows that the quickest part of the area considered is the portion of highway in the area close to the city of Bologna as expected. The second quickest part are tangential/provincial roads and statal roads, while the second slowest the urban roads. Interestingly enough, part of the highway far from the city of Bologna is considered of slowest mobility. The classification is essentially stable for all days. We here show the union of the sub-networks 4.20

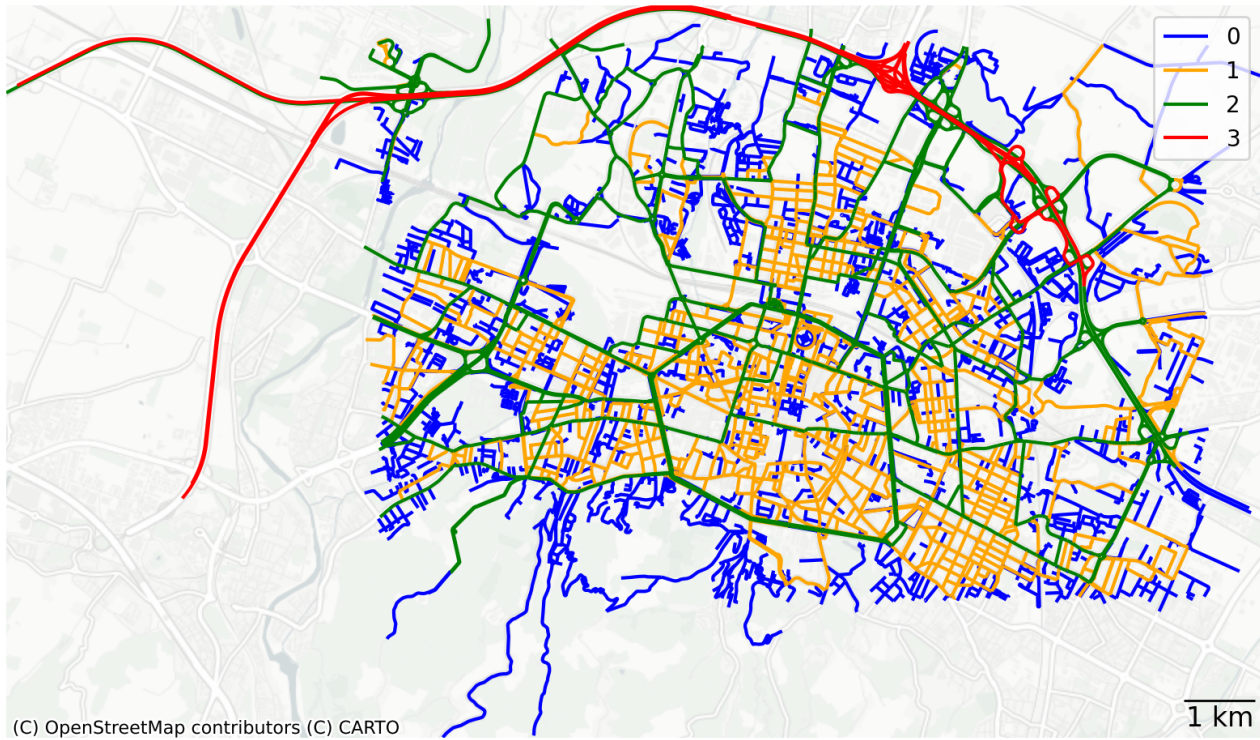


Figure 4.20: Representation of the union of the hierarchical sub-networks obtained for all different days. There is some degree of dis-homogeneity in these networks even though they highlight the fact that slow mobility is done in urban areas while high speed is concentrated along the highways both in direction of Florence and Rimini (where a small section of highway is considered to be the slowest due to shortest path reconstruction at the boundary).

4.3 Traffic phase transition

In this section, we present the results of the traffic simulator and flux reconstruction analysis introduced in Section 1.4. The data sources are detailed in Section 2.3.1, and the goals and methodologies are outlined in Section 3.3. We highlight the main features of the analysis. The focus is on the metropolitan areas of San Francisco and Rio de Janeiro.

The objective of this study is to reproduce the non-equilibrium phase transition to an absorbing state—associated with the onset of traffic gridlocks—as described in [56, 49].

We generated 40 Urban centrality indices (UCIs) for each city (see Fig. 3.11); however, simulations have only been completed for a subset of these configurations.

For Rio de Janeiro, all 40 UCIs (4 per interval of 0.1 (ΔR)) were simulated. For San Francisco, 20 UCIs (spanning the interval [0.4, 0.9]) were considered. For each (city, UCI) pair, we performed $N_{\text{ensemble}}^{\text{city,UCI}} = 8$ simulations. Each simulation corresponds to a different realization of the origin-destination (OD) assignment, varying the specific origin and destination nodes while preserving the same T_{ij} matrix on the selected grid. All measurements are computed from the unloading curve, defined as the ensemble average:

$$n_R(t) = \langle n_{R,e}(t) \rangle_{\text{ensemble}}.$$

Figures 4.21a and 4.22a display the number of individuals trapped in the absorbing (gridlocked) state at large times, $n(R, t_\infty)$, across all morphological configurations for San Francisco and Rio de Janeiro, respectively. The insets in the figures highlight the supercritical regime of $n(R, t_\infty)$. The slope of the fit in this region corresponds to the critical exponent β for each UCI. We want to highlight here the different behaviors one obtains across different morphologies that are instead presented one by one in figs. like ?? and all the others presented in A.3, where zooming in allows to get a better grasp on the goodness of fit that is not reported in terms of χ^2 but left to visualization since in most cases the fit are visibly good and already support the conclusions, and most of all stick to the procedure introduced in [99], whose observations motivated this study.

The key signature of a phase transition is the collapse of the rescaled unloading curve: plotting $n(t)t^\alpha$ at $R = R_c$ yields an approximately horizontal plateau before the second exponential drop. Above and below R_c , the curves are clearly separated, confirming the presence of a non-equilibrium phase transition. We illustrate this behavior with representative transitions for four UCIs in both cities (Figs., 4.24, 4.23). We show all the different transition in A.3.2.

Additional configurations are presented in Section A.3. It is worth noting that not all UCIs exhibit a well-defined transition and for some generated configurations we were not able to get one (since in the fixed interval $[t_{\text{start}} + t_{\text{start}_{\text{fit}}}, t_{\text{start}} + t_{\text{end}_{\text{fit}}}]$ exponential fit performs better).

The extracted β values are shown in Fig. 4.21b, 4.22b. In San Francisco, β ranges in [0.04, 2.1], whereas in Rio de Janeiro it is in the interval [0.03, 1.42]. In both cities, we observe a trend in which β increases with the absolute value of the morphological exponent α , in agreement with the findings reported in [99]. In general, smaller values of β indicate greater “resilience” of the road network to gridlock. This means that as the loading rate increases, the impact on the number of individuals becoming trapped remains relatively limited.

When comparing different cities, we observe that multiple UCIs for Rio de Janeiro yield values close to $\beta \sim 0$, while San Francisco typically shows higher values. This suggests that Rio exhibits a greater “resilience” to gridlock compared to San Francisco.

If we specifically compare the original, unmodified origin-destination setup for Rio in our simulation (see Fig. A.32c, $\beta = 0.281$) with the results from [99] ($\beta = 0.61$), we find that our simulator produces a significantly lower β . This implies improved resilience, likely attributable to the inclusion of multi-lane dynamics in our model.

In Fig. 4.25, we show $\beta(UCI)$. For San Francisco, $\beta(UCI)$ has a decreasing tendency. For Rio, this trend is less apparent. However, in both cities, large fluctuations are observed, indicating that UCI alone is not a reliable predictor of β .

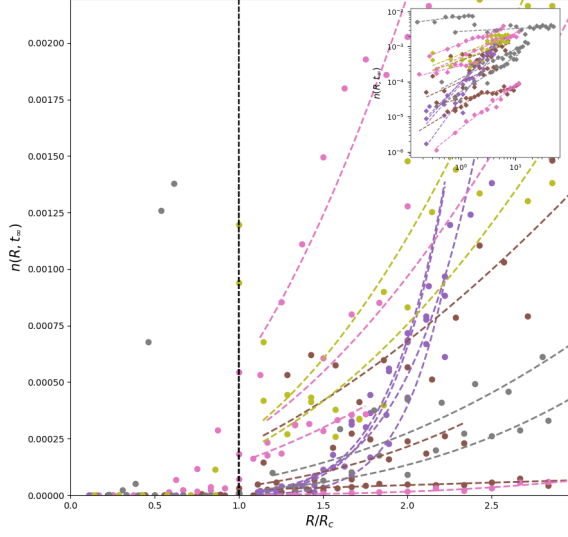
These fluctuations highlight that a network's "resilience" to congestion cannot be fully explained by the global UCI metric alone. Instead, it also depends on the structural and local properties of the road network under load.

In other words, some urban areas inherently withstand congestion more effectively under increased demand, independently of their UCI. This is evidenced by β fluctuating even across narrow UCI ranges.

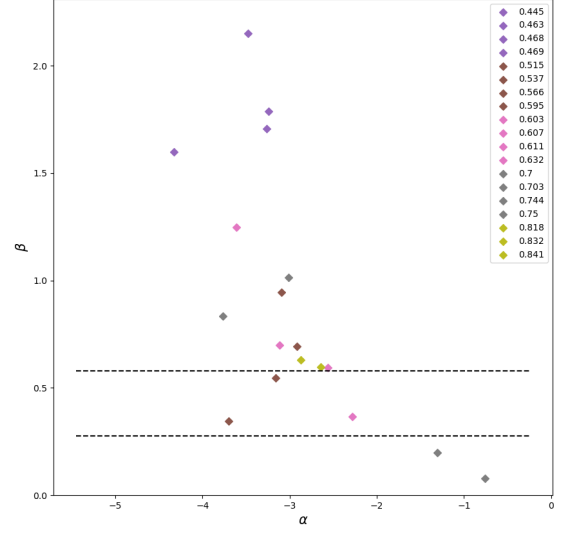
To compare how different UCIs respond to congestion, we examine the normalized delay $\frac{\tau}{\tau_c}$ as a function of $\frac{R}{R_c}$ (see Figures 4.21c and 4.22c). In both cities, we observe distinct slopes for the subcritical and supercritical regimes. Specifically, the subcritical regime exhibits a steeper increase in $\frac{\tau}{\tau_c}$, suggesting heightened sensitivity to additional load. In contrast, in the supercritical regime, the curves are less sensitive to further load.

We also analyze the dependence of $\tau - \tau_0$ on $\Gamma - \Gamma_0$ (see Fig. 4.21d, 4.22d), where τ_0 and Γ_0 denote the values below which $\tau(R)$ varies by less than 0.1, marking the free-flow regime. For San Francisco, two distinct groups of slopes emerge (with the exception of $UCI = 0.841$), whereas in Rio de Janeiro, the slopes are more uniform across most configurations, with only a few exceptions. To be noted that for the configurations studied the free flow regime consists mostly of 1 point.

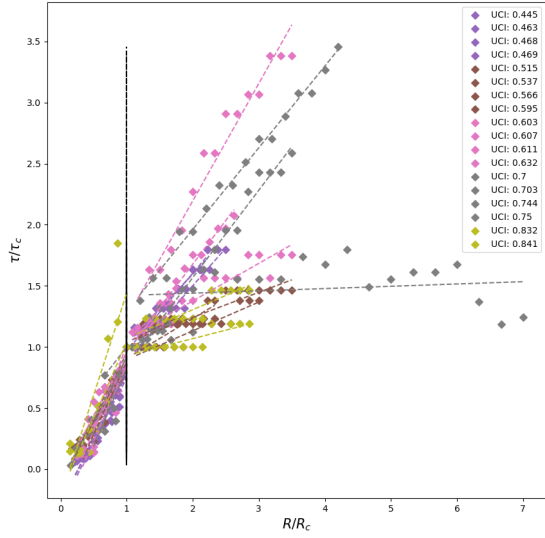
In these two case studies we already verify that the (1-2D) DP universality classes (we were looking for) are not comparable to the transition observed across all the different configurations tried, highlighting unfeasible correspondence.



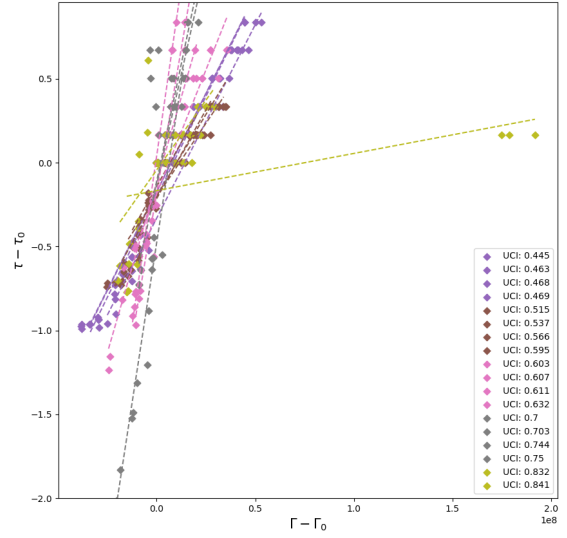
(a)



(b)

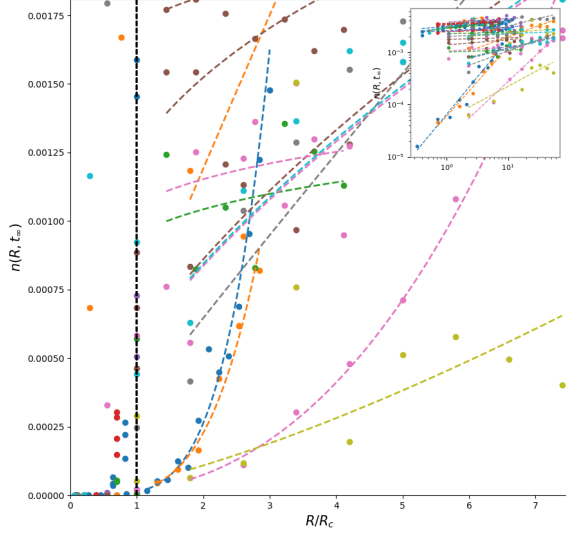


(c)

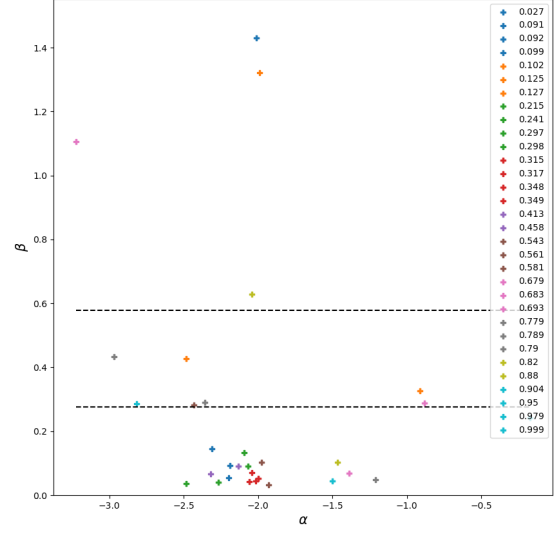


(d)

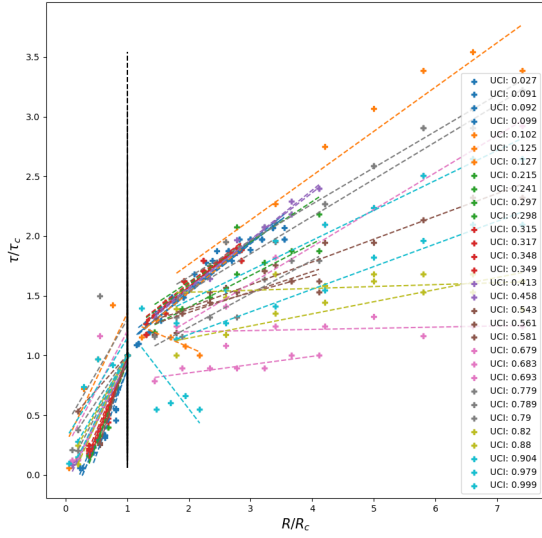
Figure 4.21: Here are represented the important features about the phase transition in San Francisco. In 4.21a is represented the number of people trapped in the road network at t_∞ for each UCI for which we have found the transition, with the fit represented in the inset. In 4.21b we see the values of β and α . We can see a broad range of critical exponents. In 4.21c we plot $\frac{\tau}{\tau_c}$ showing that the exponent tends to diverge for higher loading rates with different intensities. In 4.21d we see the linear increase of τ compared to Γ



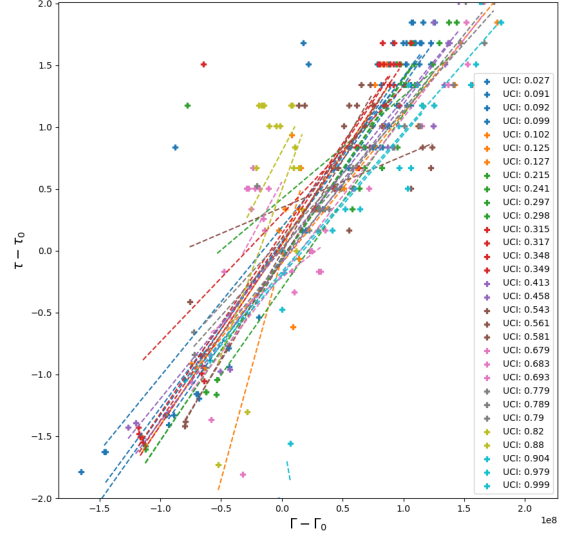
(a)



(b)

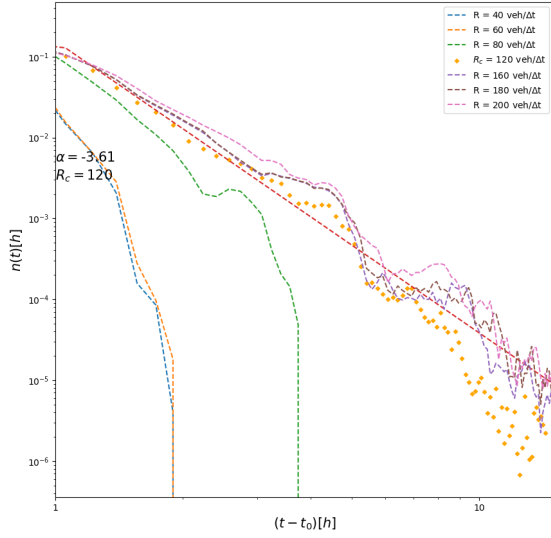


(c)

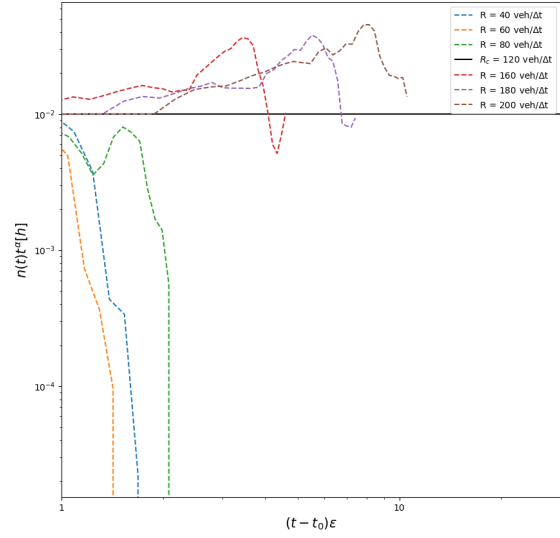


(d)

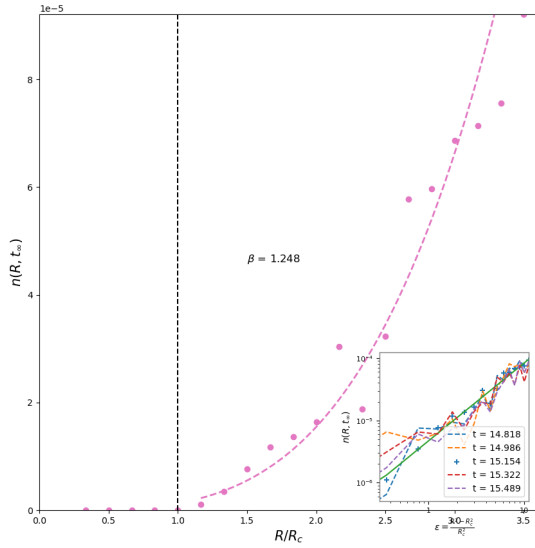
Figure 4.22: Here are represented the important features about the phase transition in Rio De Janeiro. In 4.22a is represented the number of people trapped in the road network at t_∞ for each UCI for which we have found the transition, with the fit represented in the inset. In 4.22b we see the values of β and α . We can see a broad range of critical exponents. In 4.22c we plot $\frac{\tau}{\tau_c}$ showing that the exponent tends to diverge for higher loading rates with different intensities. In 4.22d we see the linear increase of τ compared to Γ



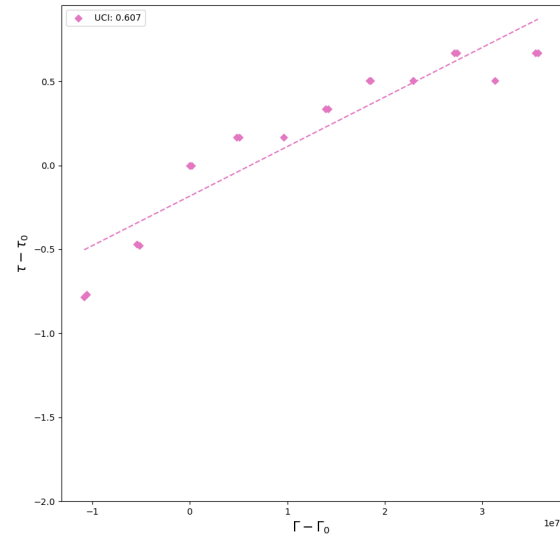
(a)



(b)

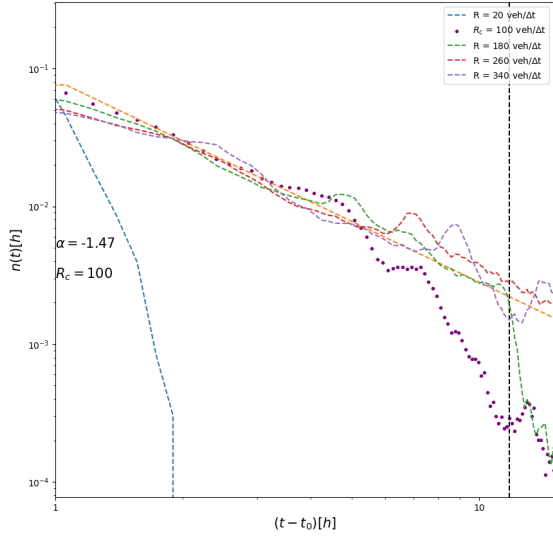


(c)

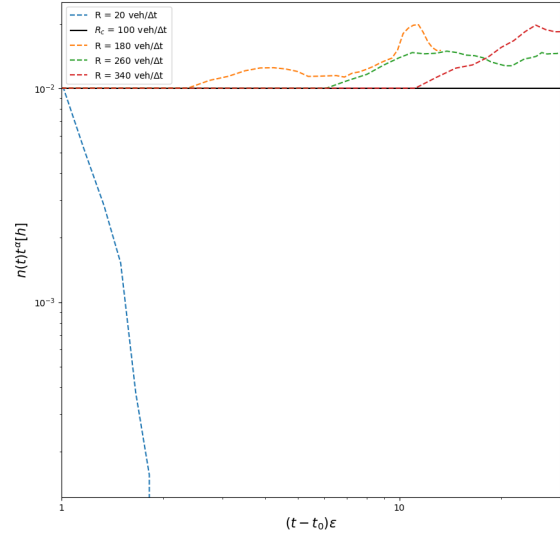


(d)

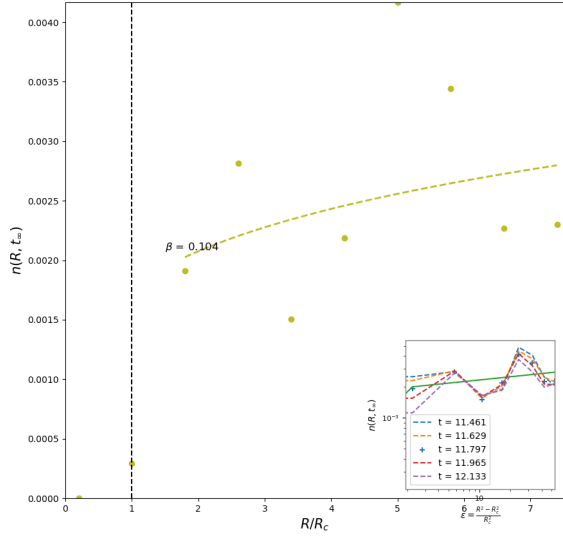
Figure 4.23: Here are represented the important features about the phase transition in San Francisco for $UCI = 0.607$ (very monocentric city). In 4.23a is represented the number of people trapped in the road network at time t , where the vertical line is t_∞ . In 4.23b we see the values $n(t)t^\alpha$ in terms of $(t - t_0)\epsilon$, where $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$. In 4.23c we plot $n(R, t_\infty)$, in the inset the fit and the lines representing $n(R, t'_\infty)$ that fit worse. In 4.23d we see the linear increase of τ compared to Γ



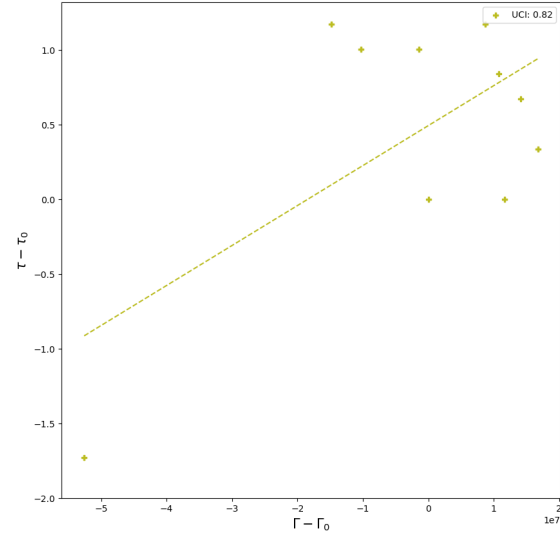
(a)



(b)



(c)



(d)

Figure 4.24: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.82$ (very monocentric city). In 4.24a is represented the number of people trapped in the road network at time t where the vertical line is t_∞ . In 4.24b we see the values $n(t)t^\alpha$ in terms of $(t - t_0)\epsilon$, where $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$. In 4.24c we plot $n(R, t_\infty)$, in the inset the fit and the lines representing $n(R, t'_\infty)$ that fit worse. In 4.24d we see the linear increase of τ compared to Γ

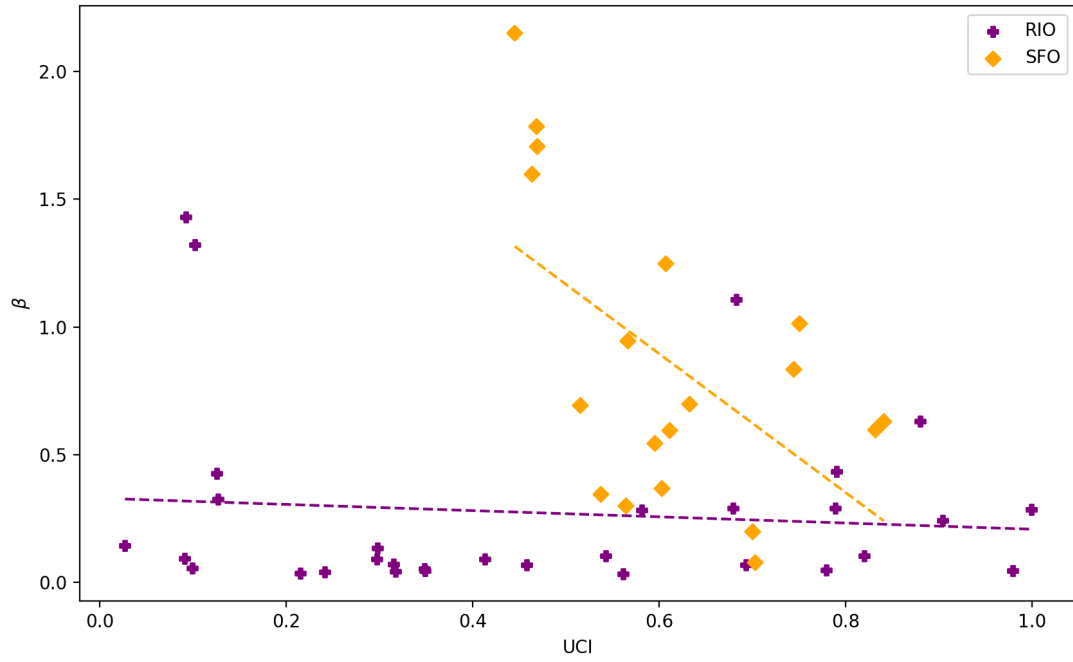


Figure 4.25: In the figure we show the dependence of β to the UCI. While San Francisco shows a tendency to diminish β increasing UCI, Rio remains flat. In either case we see fluctuations that show how UCI is not a very good indicator of the recovery of the system and we see the need of more refined descriptions.

4.4 Pedestrian mobility in Venice

In this section we present the results of the analysis of the simulator and the reconstruction of flows introduced in sect 1.5 whose data are exposed in 2.4.1 and whose goal and methodologies are listed in 3.4. The simulations were conducted just for touristic flows (i.e. *SRC_DEST* pawns). The commuters were introduced in the road network using the loading curve measured by the cameras present in the train,bus and Vaporetti station (represented in Fig. 4.29).

We highlight the draw-backs and the limitations that we were not able to overcome, while showing the partial results we can get. We highlight two different main limitations we could not overcome. On the one hand the lack of informations about vaporetti. Vaporetti (boats) is a mean of transport that is very common for tourists in Venice. It hosts a non negligible portion of tourists leaving from the train station to San Marco square. For this reason, we have included a fake generated source to the vaporetti stop. The strategy used is that a fraction of the users measured in the Scalzi station is generated there. The other limitation is the inability of describing in an accurate way the attractivity vectors of the attractions. These vectors were filled with uniform entrances of 0.3 and 0.6 (dimensionless parameters measuring the importance of the different attractions that take values $\in [0, 1]$) depending on the level of perception of importance. No further strategy was applied. The main results we have, is the identifications of directions that outgoing tourists take for two roads represented in fig. 4.28. This classification is obtained looking at the ward clustering obtained from both data and simulations (fig. 4.30) highlighting correlated groups of barriers (since aggregation time (15 minutes) is high enough to not consider retarded dynamics. The flows are represented in 4.26-4.27, where we can see similarity in flows. We stress that such a study highlights some simple behavior of touristic flows without being able to capture the overall network dynamics. The primary objective of this study was to explore how a microscopic pedestrian dynamics simulator can be employed to reproduce the spatial and temporal flow patterns measured by camera-based observations. To assess the validity of the model, we analyzed the hierarchical structure of the correlation functions of pedestrian flows (using Ward algorithms 3.4.2) across different areas. The contribution of this work to the current literature remains limited while showing the possibility to extract useful informations about most correlated roads, extracted exploiting the hierarchical structure (expected, roughly speaking, due to the different importance of paths), allowing for a framework that needs to further developed to address planning of the city for big events.

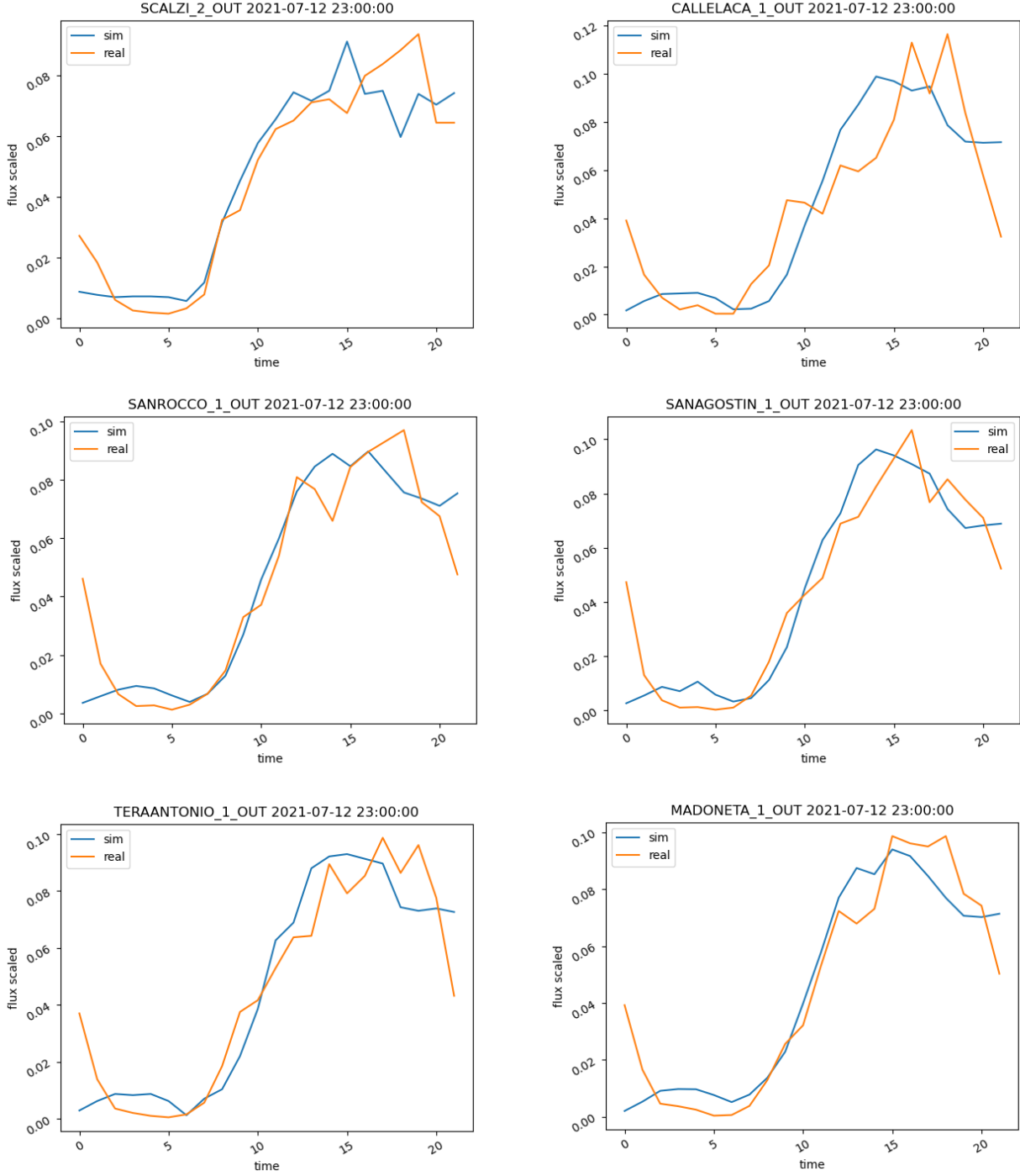


Figure 4.26: Comparison between simulation and data for out flows of people leaving measured by barriers in strada nova. We call this set of barriers "Strada nova road" [4.28](#)

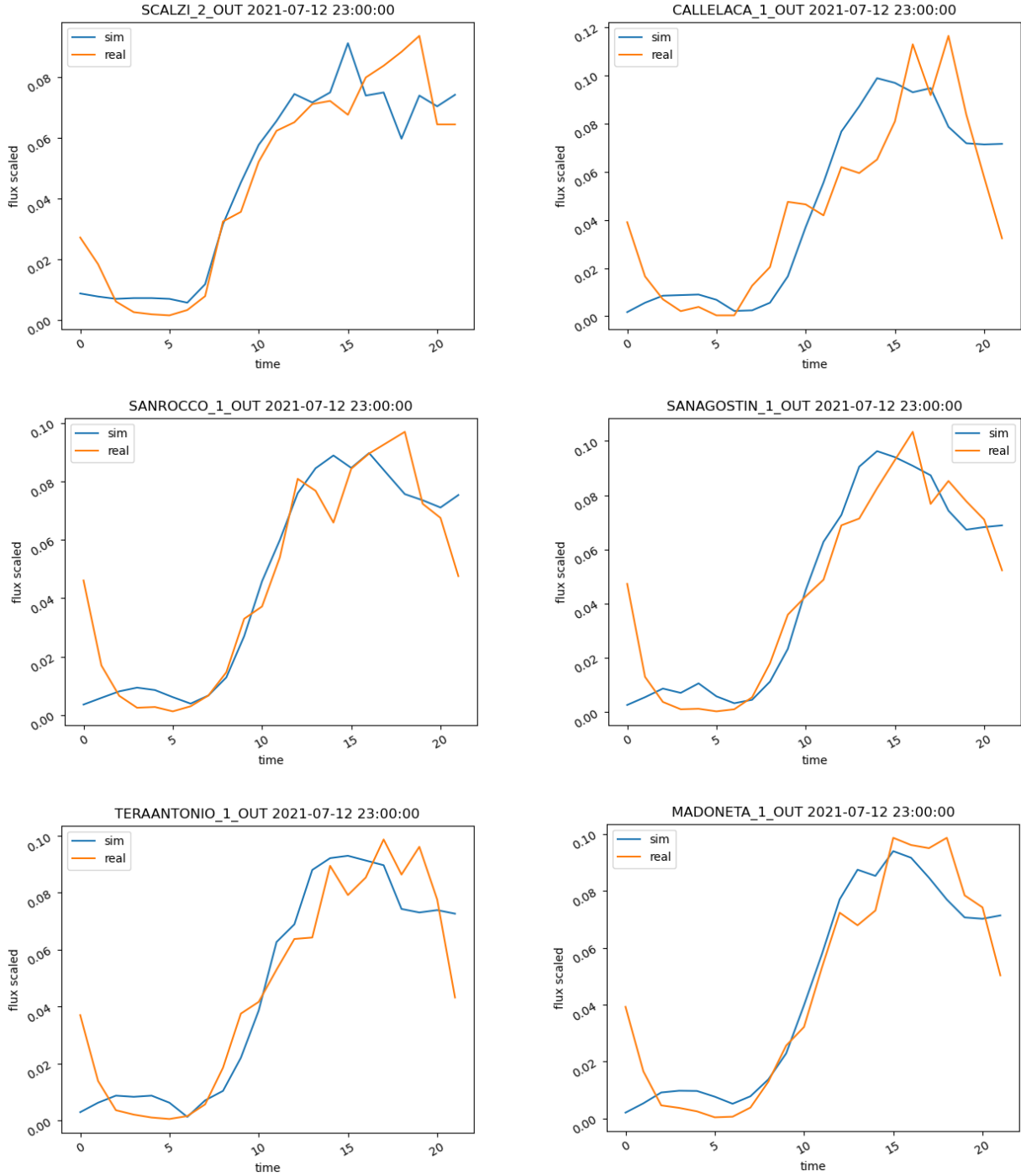


Figure 4.27: Comparison between simulation and data for out flows of people measured by barriers leaving San Marco. We call this set of barriers "San Marco road" [4.28](#)

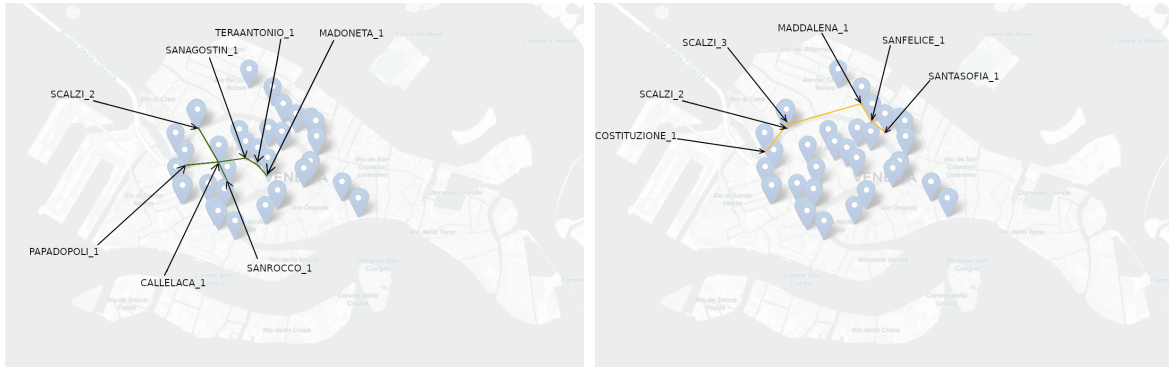


Figure 4.28: On the left the barriers associated to "San Marco road" of fig. 4.26 on the right those associated to "Strada nova road" of fig. 4.27

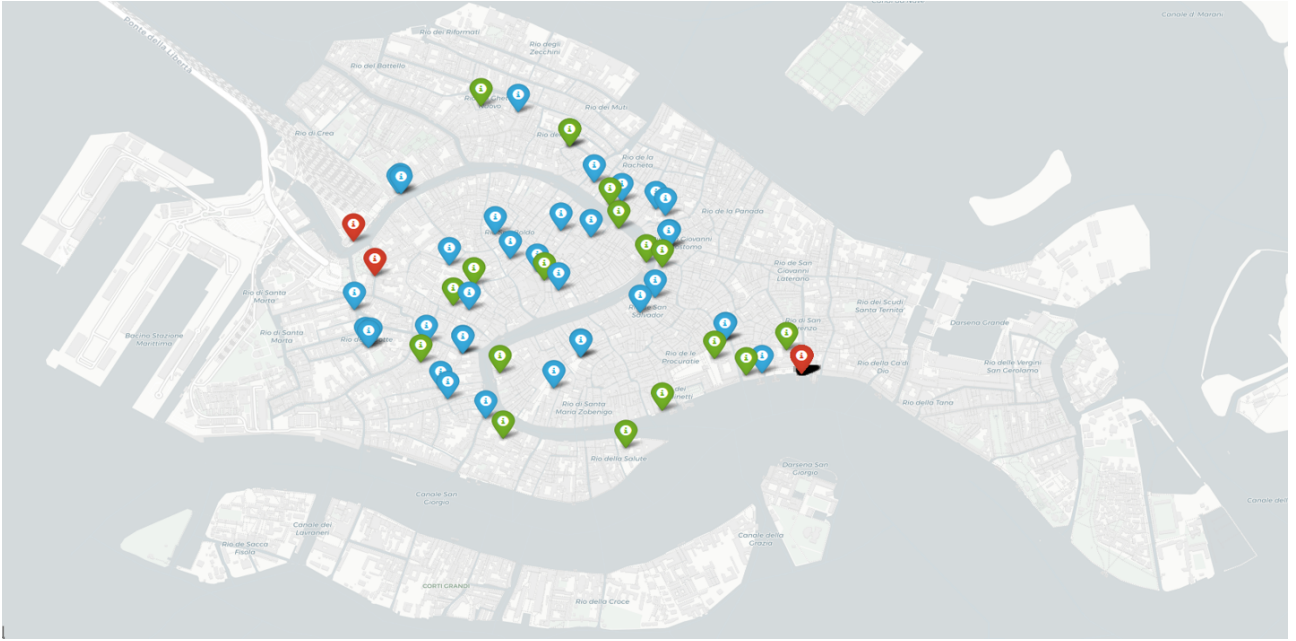


Figure 4.29: Here represented the cameras and their different functionalities roles in terms of simulation. In particular in red we have *Sources*, in blue *Barriers* and in green *Attractions*

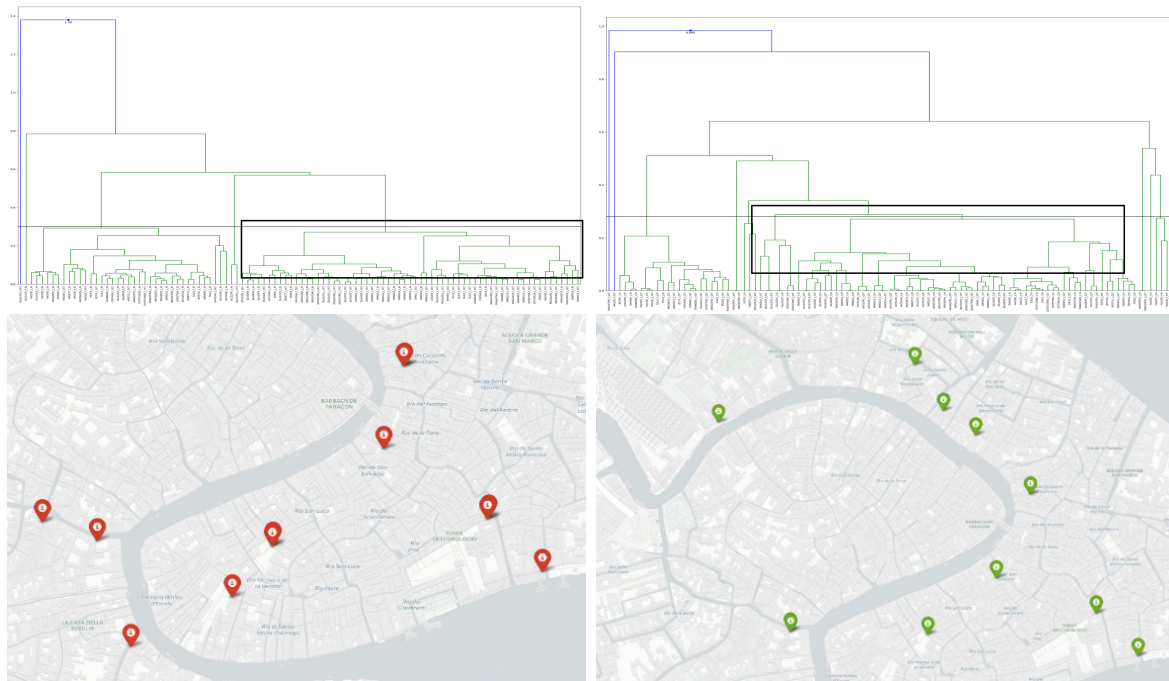


Figure 4.30: On the top are represented the dendrogram produced via Ward algorithm on the simulated data (left) for real data (right). The cut chosen a posteriori, allows to highlight clusters of cameras that behave accordingly highlighting line of flows that are present in both simulations and real data. This allows to have a rough idea of the paths that people take and are represented in 4.28

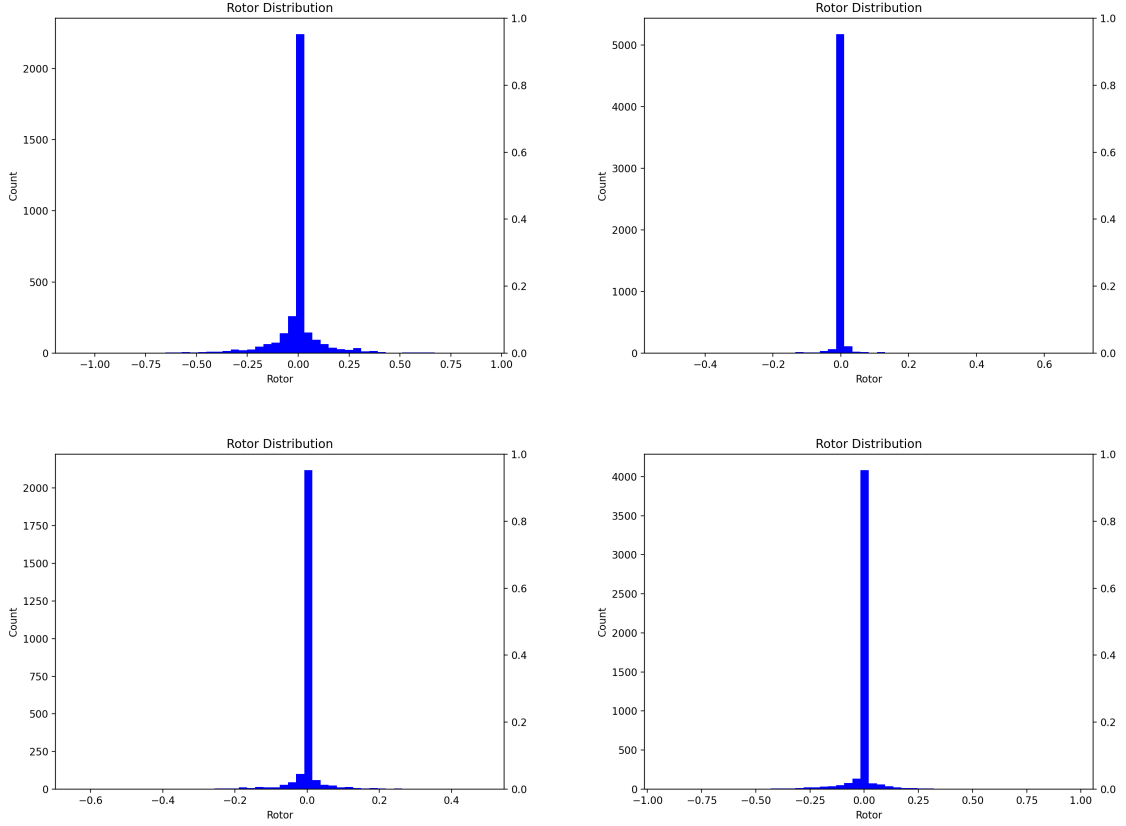


Figure 4.31: From bottom left to bottom right rotor distribution in different cells of Boston, Lisbon, Rio de Janeiro and San Francisco. This shows the local consistency of the hierarchy defined by the potential.

4.5 Field theory for Monocentrism

This section shows the initial development of the argument developed in 3.5 to define monocentrism level in 3.3 substituting in eqs. 3.19 and 3.17 the potential instead of the mass. It was developed to define polycentrism in terms of "where people prefer to go". This definition, defining an importance measure (potential), allows one to define effective preferences of people during day, therefore, capturing, effective evolution of the shape of the city, and therefore, defining more precisely the problem of 3.3, constraining it to the definition of the phase transition given the potential defined by the origin destination at peak hour (7-8) and subject to successive loads with changing potential during the day. Despite the intriguing approach, we decided not to develop further it since technical complications in the production of different configuration and the control of the related UCI. However here we show the statistical well posedness of the definitions in eqs. 3.5.4 3.31 about rotor in fig. 4.31 (local consistency) and harmonic components in fig. 4.32 and the shape of the cities in the normal morphology 4.33.

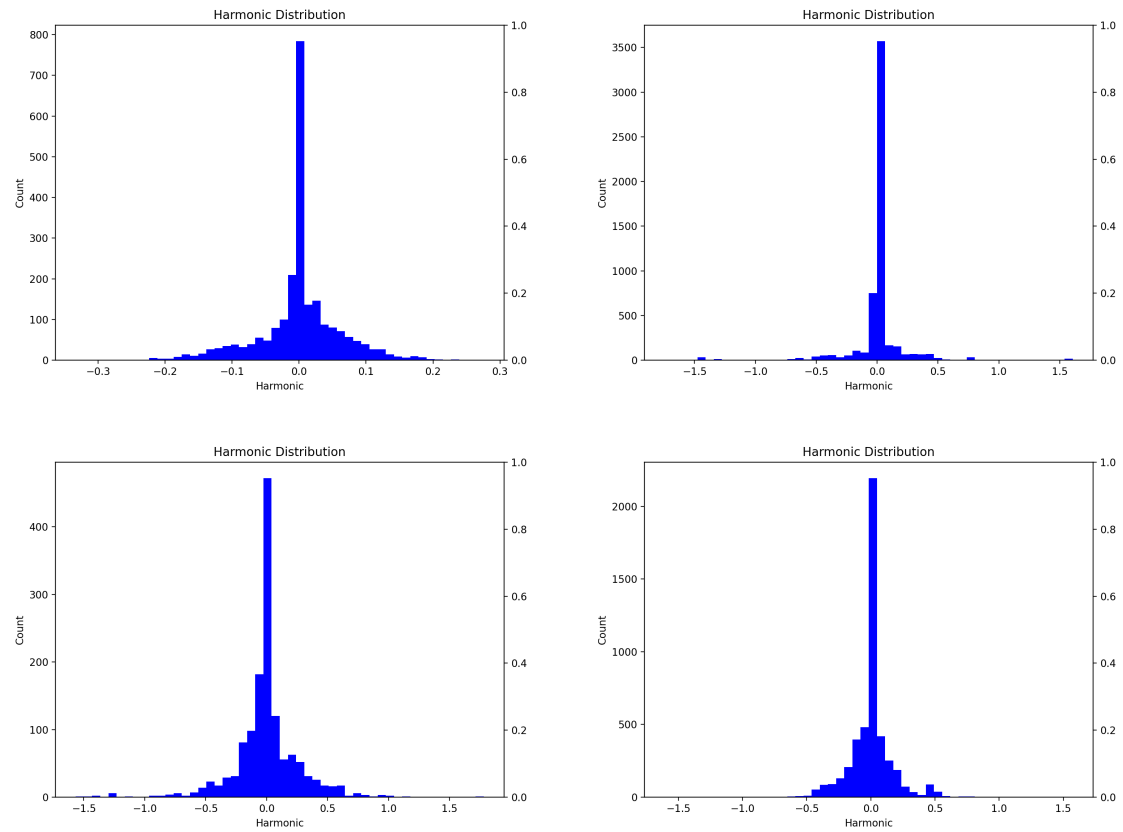


Figure 4.32: From bottom left to bottom right harmonic distribution of Boston, Lisbon, Rio de Janeiro and San Francisco. This shows the global consistency of the hierarchy defined by potential.

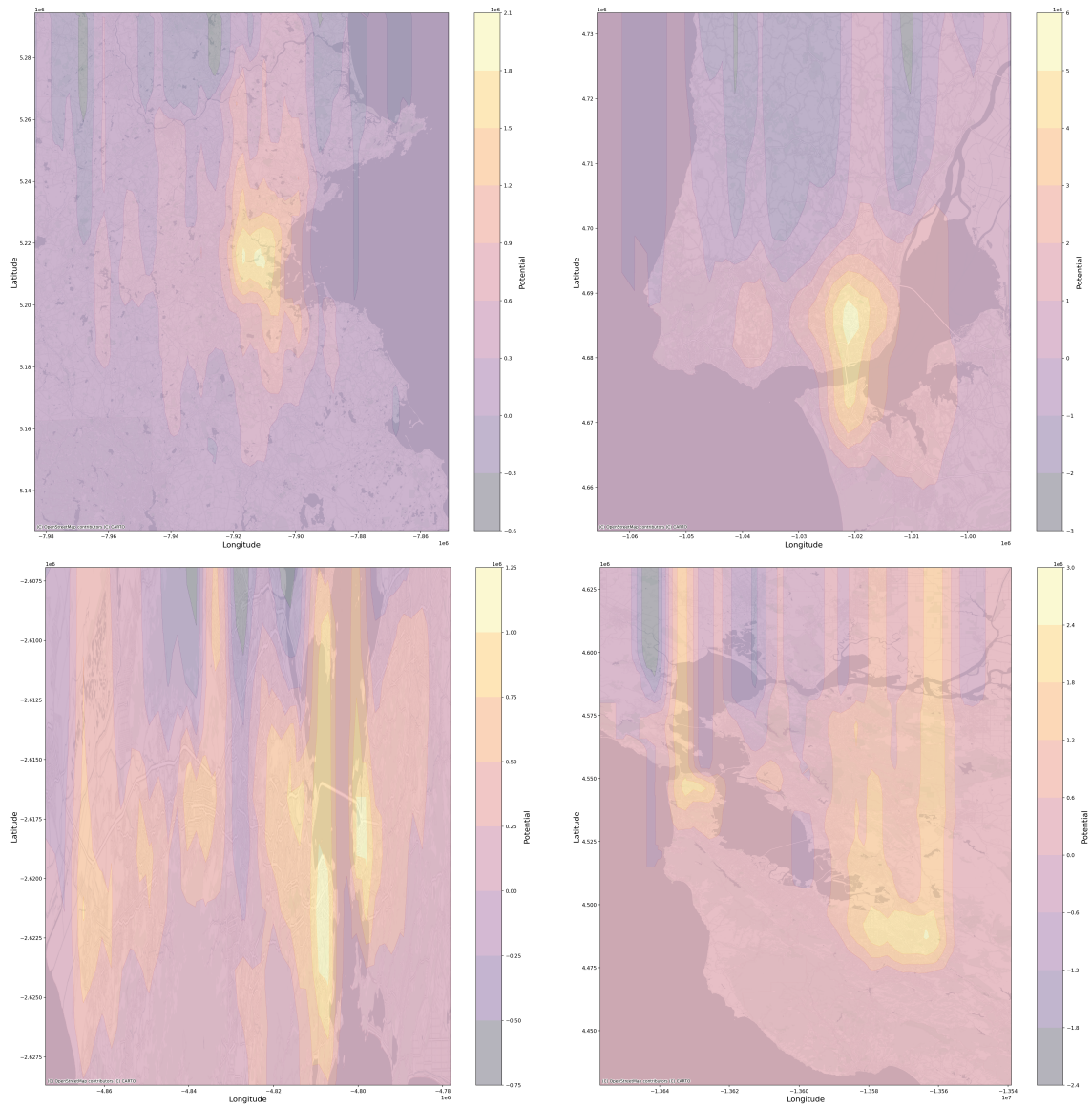


Figure 4.33: From bottom left to bottom right potential of Boston, Lisbon, Rio de Janeiro and San Francisco

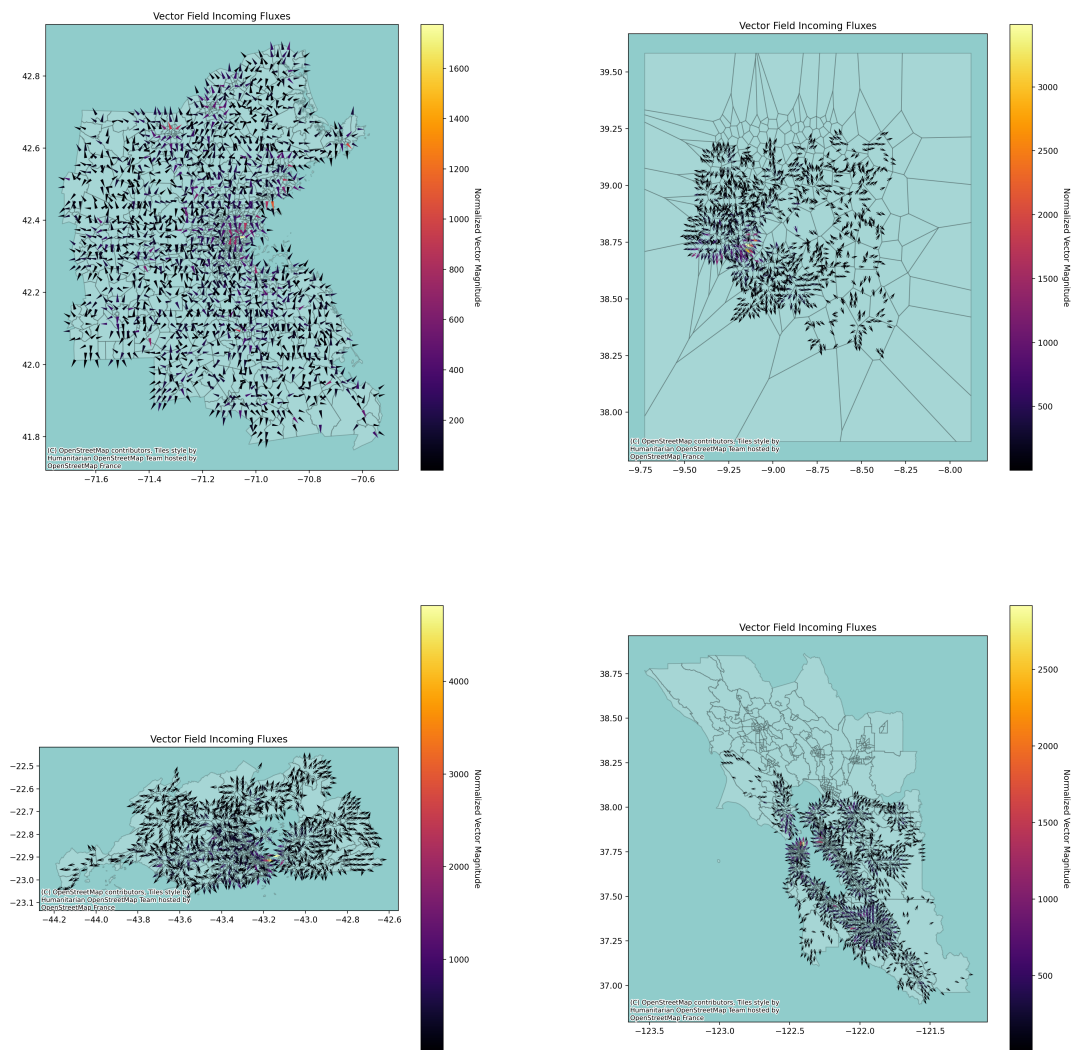


Figure 4.34: From bottom left to bottom right potential of Boston, Lisbon, Rio de Janeiro and San Francisco, is here represented the vector field of the different cities

Chapter 5

Conclusions

In this thesis we have explored different aspects of the statistical properties of human mobility at different scales. The main contribution is the characterization of out of equilibrium properties of urban systems. The general main goal has been in any case to understand how it is possible to derive the universal features for urban mobility using the statistical physics approach. Under this point of view we proposed a reductionist approach based on mesoscopic models. We have worked primarily with data, and whenever we approached models we tried to assess their validity with the available data looking at aggregated observables. We have considered both behavioral aspects of human mobility [3.1](#), studying granular data not embedded in the road network, and the interaction with the road network ([3.2-3.3-3.4](#)). Each of such studies highlights different aspects of human mobility through the same lens of statistical physics. In the same spirit of the previous chapters we divide the conclusions into multiple subsections each referring to previous chapters. We begin with the study of the quantification of change in mobility during Covid lockdown restrictions in [5.1.1](#), we continue with the characterization of the multimodality of the road network of Bologna [5.1.2](#), we have a look at a particular phase transition observed in traffic on the road network [5.1.3](#) and we finally explore the results about the simulator of pedestrian mobility [5.1.4](#) and a different pathway to define monocentrism [5.1.5](#).

5.1 Conclusion

5.1.1 Mobility changes during COVID-19

In this section, we summarize the key findings from the analysis presented in [Section 4.1](#).

We recall that the objective of this study was to describe how COVID-19 lockdowns perturbed mobility behaviors away from their equilibrium state, as introduced in [Section 3.1](#).

The main contribution of this study is to provide quantitative insights into mobility changes that can inform improved modeling of epidemic spread. Epidemic models [[6](#), [8](#)] often incorporate mobility patterns parametrically at the metapopulation level, where they define link weights in the spatial network connecting communities. These models typically rely on assumptions about baseline mobility and introduce interventions as changes in contact probabilities. We suggest that empirical information on travel behavior can support the development of more realistic models. In particular, we highlight behavioral constraints relevant for generating synthetic trajectories that can be used to estimate inter-community mixing.

The conclusions presented here are qualitatively consistent with those found in the analyses of California ([Section 4.1.2](#)) and Massachusetts ([Section 4.1.2](#)).

During the pandemic restrictions, individual mobility contracted into smaller geographic areas (see [Figures 4.8, 4.1](#)). Despite this contraction, the geometric distribution of trips remained consistent (see [Figures 4.3, 4.10](#)). Specifically, about 70% of all visited locations were concen-

trated around two main points (Figures 4.5, 4.12), which often served as bases for exploratory movements.

These exploratory trips tended to be shorter during lockdown, as indicated by changes in the tail exponent of the trip-length distribution. However, the relative frequency of long trips remained stable—likely because such trips already represented a small share of overall mobility. This suggests that a significant portion of exploratory behavior was already being fulfilled by short-range travel, which became even more dominant under restrictions.

Across various entropy measures, we observed a consistent decline in population-level information content during COVID. This reflects both a reduced number of trajectory points and a loss of structural variability in individual movement patterns. Notably, the final panel in Figures 4.7, 4.13 suggests that the entropy decline is more attributable to a decrease in the number of distinct locations than to a loss of structural complexity per se. This implies that shorter trips were not only more frequent but also more diverse, concentrating most of the residual informational richness.

By conditioning the analysis of the radius of gyration on county-level political affiliation and rurality, we obtain a more nuanced picture. While political affiliation alone does not significantly distinguish mobility levels during lockdown—likely due to the uniformity of policies—rurality does: individuals in rural areas tended to maintain higher mobility (Figure 4.6). After restrictions were lifted, however, a political divergence emerged: individuals in Democratic-leaning areas exhibited a more pronounced and persistent reduction in mobility compared to those in Republican-leaning areas.

In summary, we explored how non-pharmaceutical interventions during the COVID-19 pandemic reshaped individual mobility patterns. Our multi-faceted analysis shows that lockdowns significantly affected not only the number of trips and diversity of visited locations—consistent with prior studies such as [81]—but also the spatial structure of mobility and the intrinsic tendency to explore. These findings enhance our understanding of how extreme exogenous shocks can disrupt mobility networks, compress spatial activity, and reveal deeper behavioral dynamics related to freedom of movement and place-based exploration.

5.1.2 Characterization of the multimodal structure of the road network in Bologna

The goal of this study was to measure traffic from GPS data exploiting the functional properties of the road network, and characterize statistical properties of paths associated to mobile phone trajectories. In conclusion, we have characterized the functional partition of mobility in the city of Bologna 4.20. Specifically, we observed that people’s travel behavior, measured in terms of time and path length, remains consistent across different days, as shown in Fig. 4.14. Aggregate mobility exhibits temporal heterogeneity, as indicated by the power-law behavior fig. 4.14c, while the length distribution suggests a characteristic urban travel distance of approximately 4 km, as evidenced by the exponential fit fig. 4.14a. The bump in the tail of the distribution suggests the presence of multiple mobility classes for longer trips, where travelers transition between different layers of the network [103]. The disaggregation of these distributions reveals power-law scaling for travel times (fig. 4.14d) in medium-length mobility classes, while short-range mobility follows an exponential distribution. Combined with the observed mixture of speeds within individual trips (Fig. A.9), this suggests that users struggle to optimize their travel time budgets due to traffic conditions, which force them into longer trips.

Moreover, we find that characteristic trip lengths increase linearly across mobility classes. The classification of mobility patterns proves particularly valuable for detecting traffic signals in these datasets. While traditional Macroscopic Fundamental Diagram (MFD) measures (Fig. 4.16) do not effectively capture traffic conditions, our approach successfully identifies deviations

from free-flow conditions (Fig. 4.18). This new measure allows for the analysis of traffic patterns in different subnetworks separately, enabling comparisons of traffic severity across days (Fig. 4.19). Notably, during holiday periods (December 30–January 1), traffic congestion is less severe than on regular days, particularly in the fastest mobility layer (highways and the Tangenziale). In Fig. 4.19, we observe that the congestion degree is particularly high for classes 1 and 2, which correspond to a spatially limited yet widespread pattern of mobility within the urban area. This suggests that, in order to reduce congestion in Bologna, it would be necessary to promote a shift in citizens’ travel habits—specifically, by encouraging the use of collective transport modes such as trams. Conversely, the ring road exhibits localized congestion primarily during peak hours, and a potential expansion (e.g., doubling) of the ring road would not provide a structural solution to the broader mobility issues faced by the city. This final result underscores the strength of this data-driven approach in capturing urban traffic patterns. Such a result is of interest in the context of digital twin [59].

5.1.3 Traffic Phase Transition

In this section, we summarize the main findings presented in Section 4.3.

The primary objective of this study was to investigate the relationship between urban morphology—quantified as described in Section 3.3.3—and the universality class of directed percolation, as discussed in Section 3.3.4 and previously reported in [99].

To this end, we analyzed five cities, considering both their urban cores and surrounding extra-urban areas (see Figure 2.3). Based on available origin-destination data, we calibrated a gravity model (Table 3.2, Fig. 3.10) and generated synthetic mobility flows by modifying the population distribution, following the procedure described in Section 3.3.6. This allowed us to produce 40 distinct mobility configurations per city, each associated with a different Urban Centrality Index (UCI), as shown in Figure 3.11.

These configurations were used to simulate microscopic traffic dynamics, focusing on the emergence of a non-equilibrium phase transition to an absorbing state—characterized by persistent congestion—consistent with the findings of [99, 49]. Our analysis aimed to understand how urban morphology shapes the onset and nature of this transition.

We completed the full analysis for Rio de Janeiro and a partial analysis for San Francisco (20 out of 40 configurations). The simulation results demonstrate that the critical behavior observed does not consistently align with the directed percolation universality class. Specifically, the critical exponent β varies significantly across configurations, ranging from 0.03 to 2.1 (1.4) in San Francisco (Rio). This discrepancy may stem from the more realistic multi-lane dynamics incorporated in our simulator, as discussed in [121].

Moreover, while San Francisco shows a decreasing tendency in $\beta(UCI)$, suggesting that more monocentric configurations tend to trap fewer individuals, this trend is not clearly observed in Rio de Janeiro. In both cities, large fluctuations in β are evident even for UCIs differing by less than 0.1 (Fig. 4.25), indicating that UCI alone is not a reliable predictor of network resilience or congestion onset.

The behavior of the unloading curves near the critical load R_c confirms the presence of a continuous phase transition in most configurations, characterized by the emergence of a plateau in the rescaled quantity $n(t)t^\alpha$ and subsequent exponential decay. Additional insights were obtained by analyzing the normalized delay τ/τ_c as a function of R/R_c (Figs. 4.21c, 4.22c). In both cities, the subcritical regime is more sensitive to increases in load, while the supercritical regime shows a more stable behavior. The analysis of $\tau - \tau_0$ versus $\Gamma - \Gamma_0$ (Figs. 4.21d, 4.22d) shows the increasing unloading time in terms of unloading rate: San Francisco exhibits heterogeneous responses across UCIs, while Rio shows more uniform behavior.

In summary, our results indicate that the resilience of urban road networks to congestion cannot be fully captured by global morphological metrics such as the UCI. Instead, resilience appears to emerge from a complex interplay between local network structure, microscopic traffic dynamics, and the spatial distribution of demand. While UCI provides a coarse measure of monocentricity, it fails to account for local bottlenecks, lane-level structure, and interaction dynamics, all of which significantly influence the critical behavior of the system. In these two case studies we already verify that the (1-2D) DP universality classes (we were looking for) are not comparable to the transition observed across all the different configurations tried, highlighting unplausible correspondence.

5.1.4 Pedestrian mobility in Venice

The goal of this study was to explore possible configurations of the simulator and develop a pipeline to validate the simulator of pedestrian mobility of tourists described in section 3.4. For this reason we have developed a framework to launch different simulations considering different parameters and different configurations of sources to capture the touristic flows in the city of

Venice. We have compared real data flows measurements with simulated flows in 53 different positions 2.4.1 via the cluster of mobility obtained from Ward clustering 3.4 applied to the time serie of the flows collected every 15 minutes. We remark the known limitations of our method being On the one hand the lack of informations about vaporette that is a mean of transport that is very common for tourists in Venice and the lack of an accurate way to describe the attractivity vectors of the attractions. The results are pretty unsatisfactory since just a small subset of flows can be reconstructed with good quality like the ones measured in the road of "San Marco" and "Strada Nova" as in fig. 4.26, 4.27. In these cases we find the flows correlating and the ward algorithm highlighting the similarity fig. 3.4.2. From these results we can just infer two qualitative directions that people choose while building their trips from the train and bus station fig. 4.28. The contribution of this work to the current literature remains limited while allowing for a framework that needs to be further developed to address planning of the city for big events.

5.1.5 Another Method to Measure Monocentrism

The contribution of the methods introduced in Section 3.5 and the related results in Section 4.5 is the formulation of a mathematically rigorous measure of hierarchy derived from origin-destination (OD) data via the construction of a potential field. This potential can be employed to define levels of monocentrism in Section 3.3 by replacing population mass in Equations 3.19 and 3.17 with values derived from the potential field.

This approach was developed to describe polycentrism in terms of "where people prefer to go" rather than "where people reside." By doing so, it captures the effective preferences of individuals throughout the day, reflecting the dynamic evolution of urban structure and offering a more elastic and temporally sensitive definition of monocentricity. In this way, cities are interpreted not as static entities but as evolving systems shaped by daily mobility patterns.

The method is conceptually dual to the hotspot analysis proposed by [77], but with a focus on flows. It exploits the irrotational nature of the mobility-derived vector field first introduced in [85], connecting it to tools from statistical Hodge theory and combinatorial hierarchy frameworks, which allow for a rigorous mathematical characterization of urban centrality [9].

Despite the elegance of the framework, we opted not to pursue it further in this work due to technical challenges in its application. In particular, generating synthetic configurations of the Urban Centrality Index (UCI) using the potential-based approach encountered ergodicity issues not present in population-based models. Nevertheless, the underlying mathematical framework remains internally consistent and conceptually robust.

Beyond its technical challenges, this approach offers a novel perspective on urban centrality: one that prioritizes mobility and functional centrality over static population distributions, thus opening the door to a temporally dynamic understanding of urban structure.

Appendix A

Appendix

We here report additional material related to the studies proposed in the main text.

A.1 Mobility changes during COVID-19

Table A.1

	15 January - 15 March	15 March - 15 May	15 May - September
15 January - 15 March	1	0.6	0.6
15 March - 15 May		1	0.4
15 May - September			1

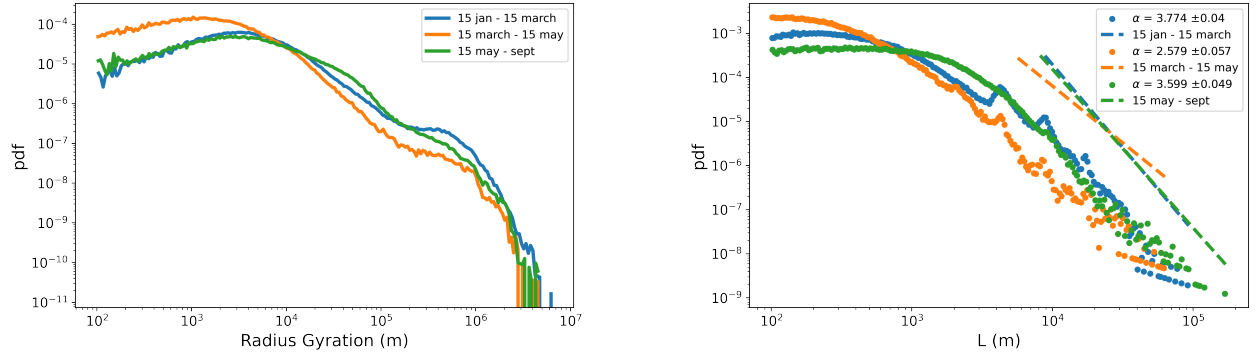


Figure A.1: On the left: Distribution of radius of gyration among in three different periods. On the right: Distribution of distances among successive stop' points in three different periods and their corrective power laws Time threshold = 8 h

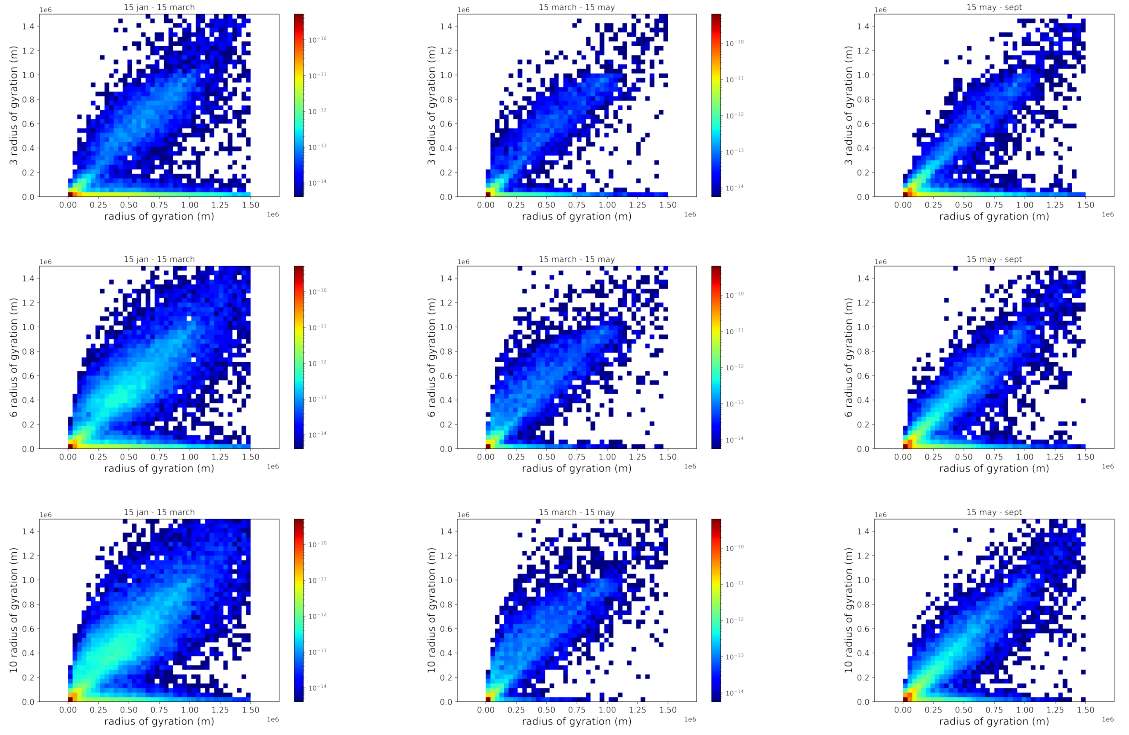


Figure A.2: Distribution of (radius of gyration, k-radius of gyration) in three different periods. From top to bottom $k = 3, 6, 10$. From left to right period of measure (15 January - 15 March, 15 March - 15 May, 15 May - September). Time threshold = 8 hours

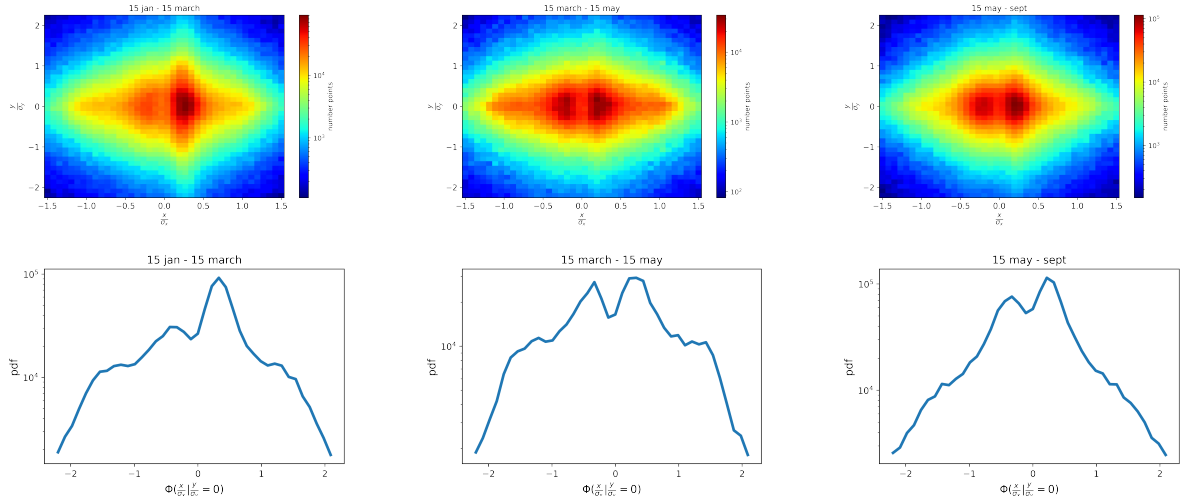


Figure A.3: Distribution of positions of people stops' points in their intrinsic reference frames re-scaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). Bottom figures are the conditional probability fixing the coordinate $y = 0$. Case of time threshold of 8 hours.

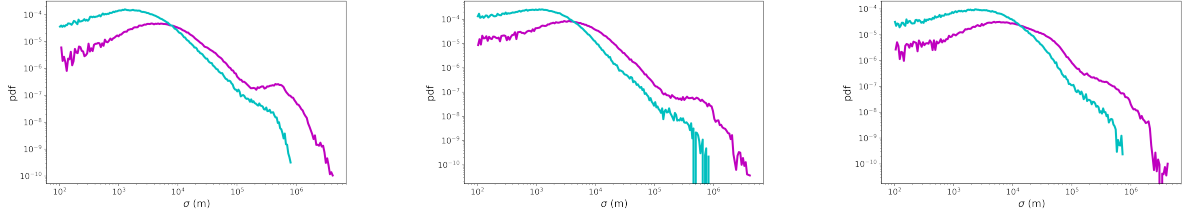


Figure A.4: Distribution of variances of trajectories stops' points in their intrinsic reference frames re-scaled by the variance of the rotated coordinates in the x,y-rotated axes. From left to right, different periods (15 January - 15 March, 15 March - 15 May, 15 May - September). time threshold of 8 hours.

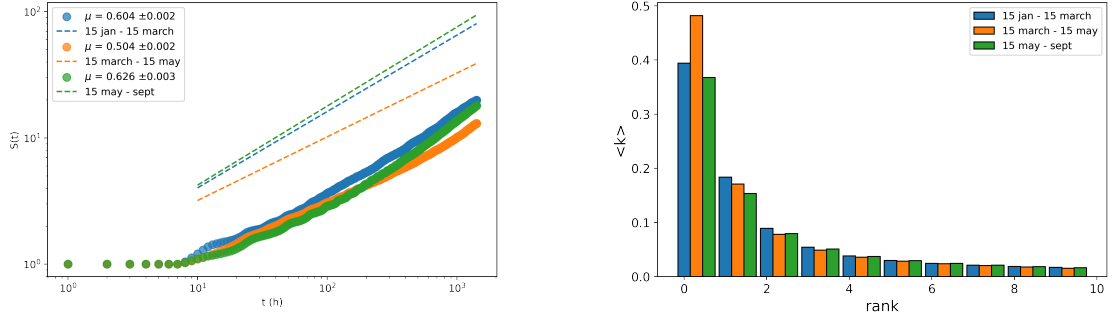


Figure A.5: On the left: Average number of stop points visited in time. On the right: Distribution of frequency of visits of first ten preferred places for each user. Case of time threshold of 8 hours.

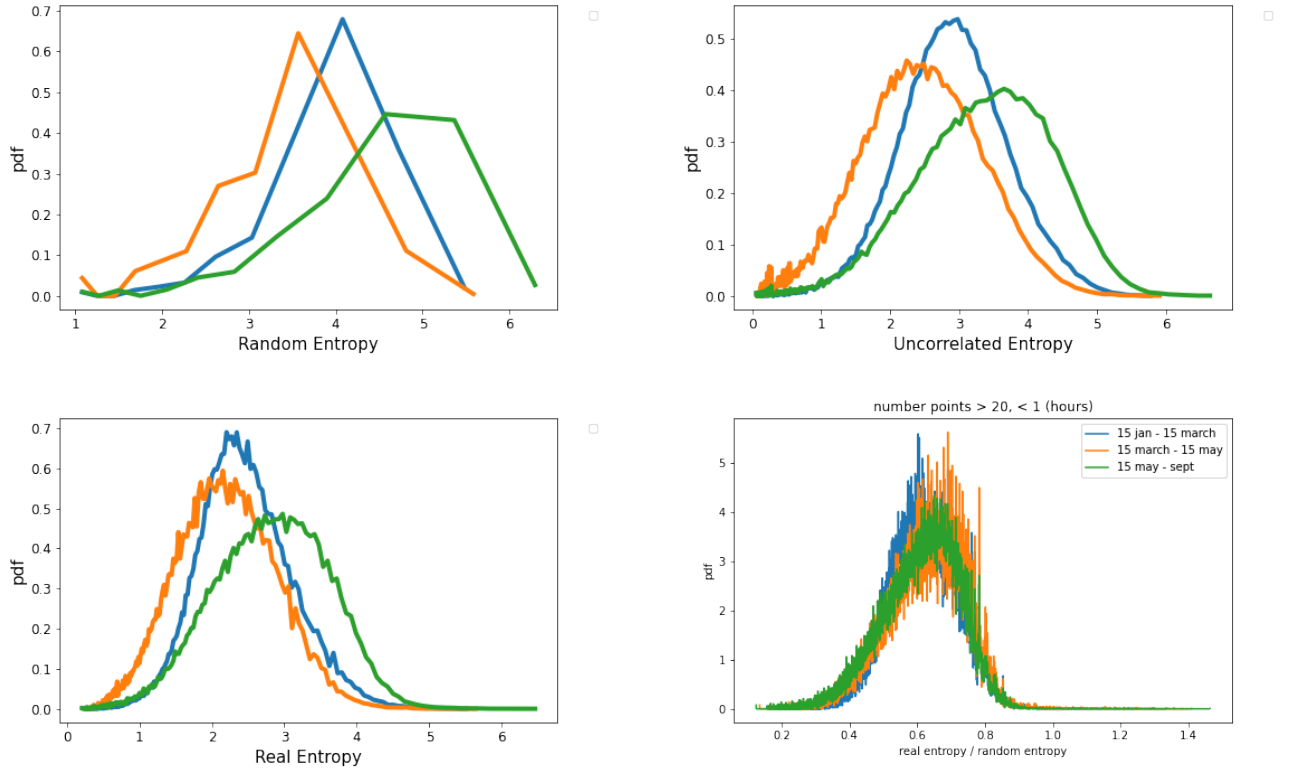


Figure A.6: Distribution (from left to right) of random, uncorrelated and real entropy. On the bottom: compression rate Time threshold of 8 hours

A.2 Characterization of the multimodal structure of the road network

We have shown in the main text material about averaged distributions and expectation values over days. We put here some additional material that covers results of different days separately.

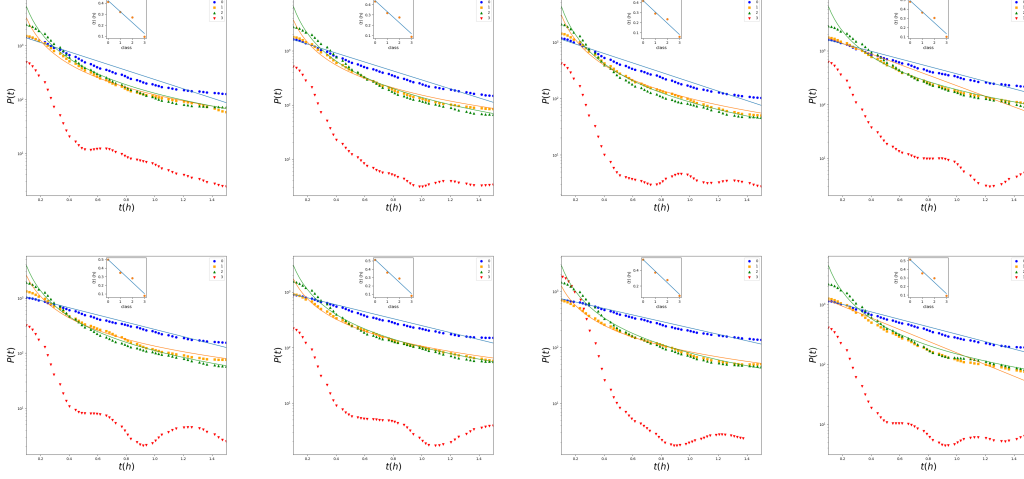


Figure A.7: Distribution of the times for paths for different days in the dataset studied in log log scale. The slowest class (0) is consistently fitted in exponential across different days. This distribution is in accordance with the energy hypothesis for travelling while the power law distributions mix different classes of travels as can be seen from [A.9](#). From top to bottom, left to right the dates follow the order: 2022-01-31, 2022-07-01, 2022-08-05, 2022-11-11, 2022-12-30, 2022-12-31, 2023-01-01, 2023-03-18. The inset represents the average time of travel per each class with relative error (std). The average trip time, decreases linearly and in particular the average time for the last class is around 6 minutes which means that are captured very short trips. The lack of fit for the quickest class is not shown as it contains too few paths and do not fit well.

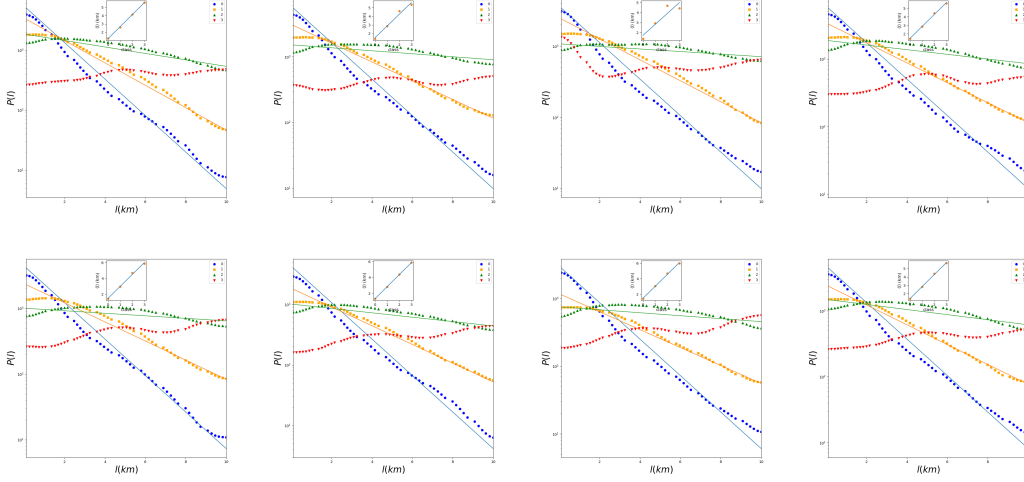


Figure A.8: Distribution of the lengths for paths in different classes for different days in the dataset studied in log log scale. All classes described here show exponential fitting with increasing average for each class. The fastest class is not fitted as contains too few data and do not fit well. From top to bottom, left to right the dates follow the order: 2022-01-31,2022-07-01,2022-08-05,2022-11-11,2022-12-30,2022-12-31,2023-01-01,2023-03-18. Values of the fit are contained in tab. A.3

Class	2022-01-31	2022-07-01	2022-08-05	2022-11-11	2022-12-30	2022-12-31	2023-03-18
1	-1.6859	-1.6500	-1.7435		-1.5269	-1.4244	-1.5080
2	-1.9014	-2.0035	-1.9649	-1.8349	-1.8126	-1.7317	-1.7402

Table A.2: Values of power law exponents for different dates and classes of mobility. $P_{day}^{class}(t) \propto e^{-\beta t}$

Class	2022-01-31	2022-07-01	2022-08-05	2022-11-11	2022-12-30	2022-12-31	2023-03-18
0	-0.5051	-0.5006	-0.4248	-0.4525	-0.4108	-0.4232	-0.4417
1	-0.4005	-0.3540	-0.3392	-0.3008	-0.3410	-0.3303	-0.3375
2	-0.2722	-0.2265	-0.2136	-0.2357	-0.2098	-0.1914	-0.1991
3	-0.1855	-0.2019	-0.1303	-0.1798	-0.1303	-0.1603	-0.1534

Table A.3: Values of exponents for different days $P_{day}^{class}(l) \propto e^{-\lambda l}$

In the end we give an idea on the resolution in time of successive points in a trajectory. We see from fig.

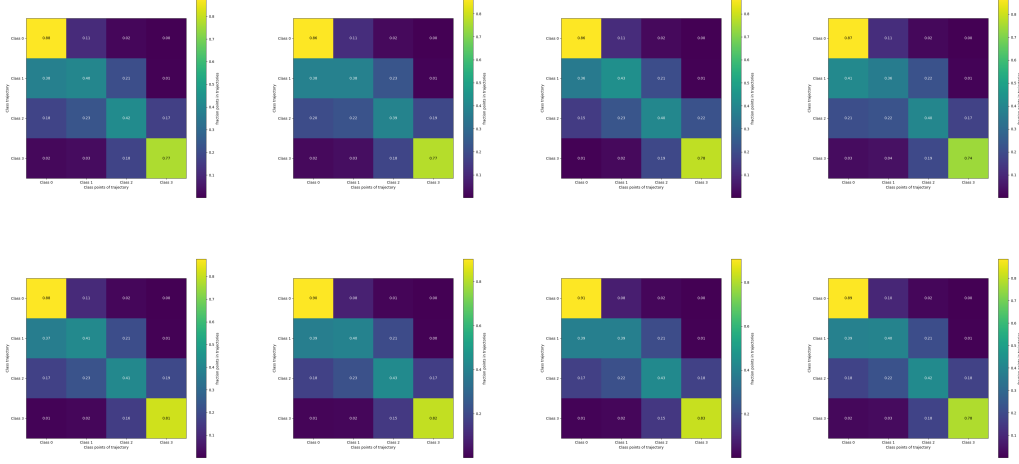


Figure A.9: Distribution in different classes of points in the trajectories. We can see that trajectories of the classes 1 and 2 contain points in different layers of mobility showing therefore heterogeneity in them. This fact can be interpreted as the difficulty of travellers to estimate travel times and preserve energy principle in [72]

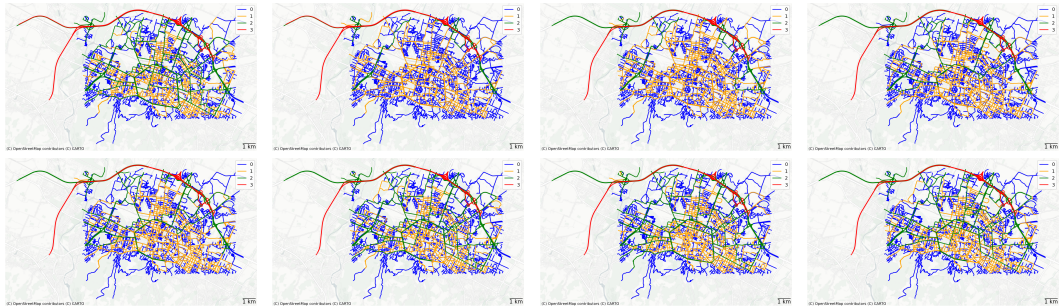


Figure A.10: Representation of the sub-networks in Bologna for different days in the dataset for all different days from left top to right bottom: 2022-01-31, 2022-07-01, 2022-08-05, 2022-11-11, 2022-12-30, 2022-12-31, 2023-01-01, 2023-03-18. We have problems of classification for the slowest speed class at the boundary where the highways are connected are the shortest path to connect outskirts and insides of the urban area of Bologna.

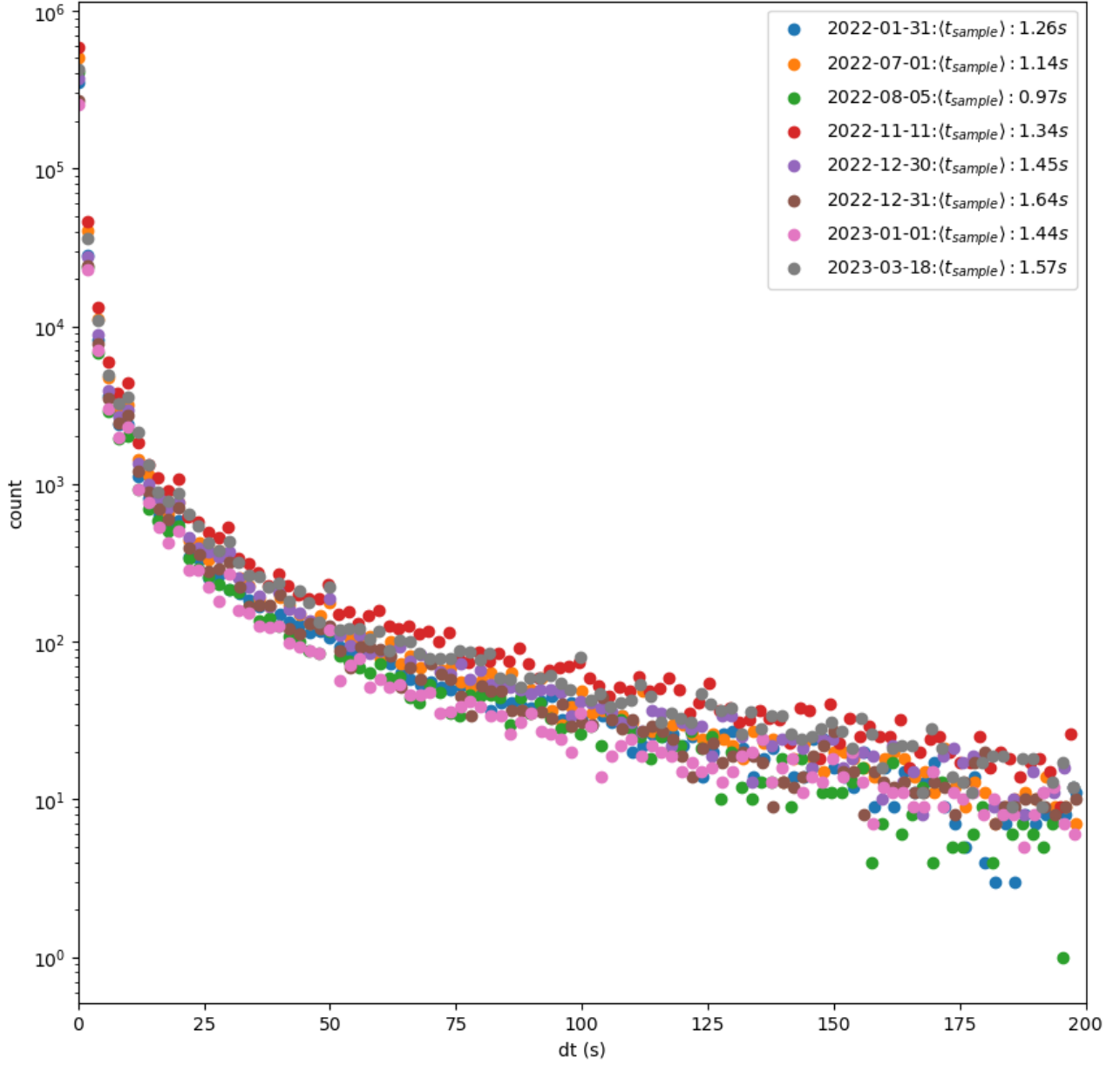


Figure A.11: Distribution time intervals between two successive points in a path. The distribution preserves a fat tail and therefore highlighting heterogeneity in the trajectories at hand, however, we see that the average interval between two successive points is less than 2 seconds consistently for all days, granting a good granularity of data

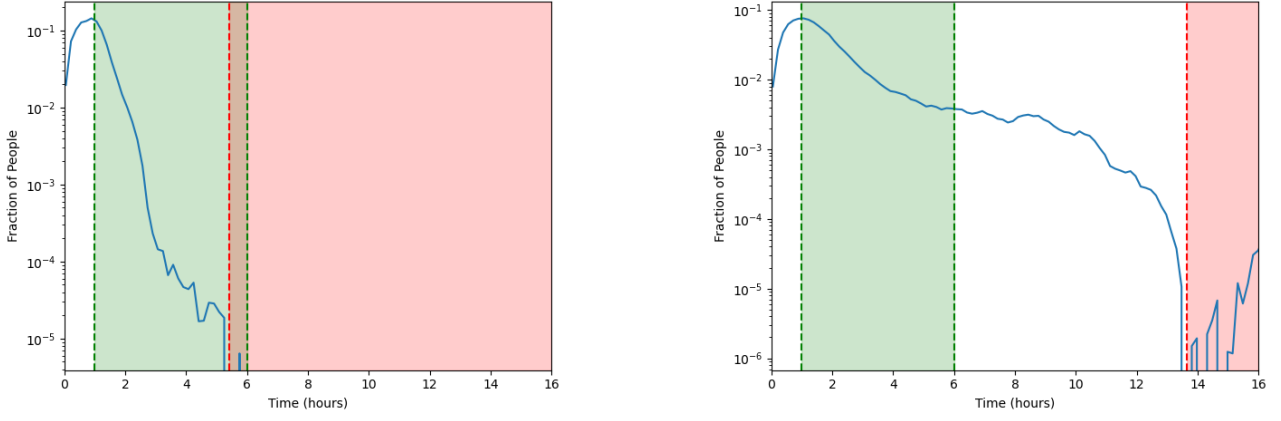


Figure A.12: In either cases the portion of the curve we fit is just the portion in the green area. On the left $t_{n_t=0} < t_{start} + t_{end_{fit}}$, therefore we consider noise the queues formed on small number of nodes. On the right $t_{n_t=0} > t_{start} + t_{end_{fit}}$ and therefore we fit all the interval cutting a portion between 6 and 10 hours after t_{start}

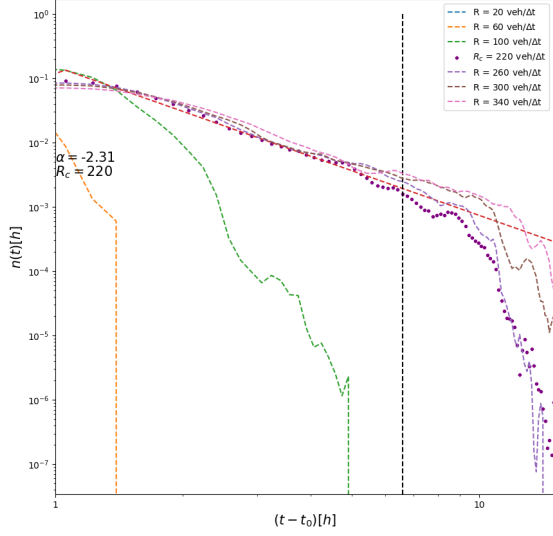
A.3 Traffic phase transition

A.3.1 Fitting procedure

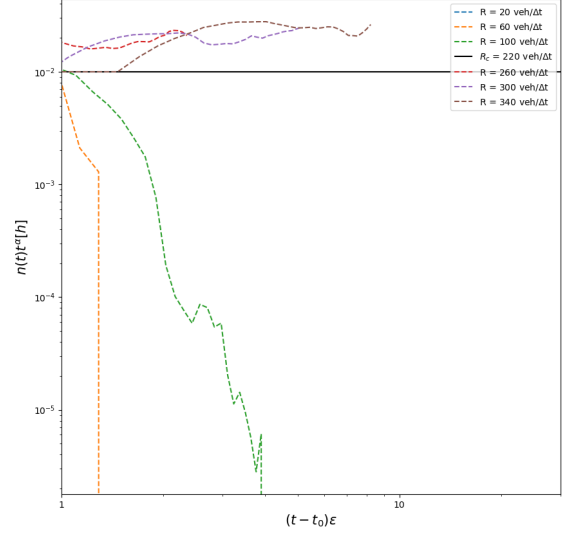
We here specify the choices of the fitting procedures for both the choice of the exponents α and β . The computation of α for a single R of a given city of defined UCI is done as a linear fit on log-x and log-log scaled values of $n(t)$ and t and in the interval of time filtered by the rules below. Let $t_{start_{fit}} = 1hour$ and $t_{end_{fit}} = 6hours$, we compute the fit on the curve $n(t)$ s.t. $t \in [t_{start} + t_{start_{fit}}, t_{start} + t_{end_{fit}}]$ and we call this procedure $filter_0$. As we are fitting in log transformed time series we need to filter 0s. In particular, let $t_{n_t=0}$ be the time in which we first measure $n(t) = 0$, we cut $n(t)$ s.t. $t < t_{n_t=0}$ ($filter_1$). The principle is represented in fig. A.12. $filter_0$ is applied since the curve unload curve $n(t)$ presents an exponential drop in the first hour that dominates the fit. $filter_1$ is needed to being able to compute the log-log fit without need to interpolate the missing values. We stress that the error of the fit is computed on the selected curve after $filter_0$ and $filter_1$ are applied. As for β once we find R_c we consider all such $R \geq R_c$ and collect $n(R, t)$ for $t > t_\beta = 7hours$. The fit for β is computed just for the cases for which we have at least 4 points above the critical point (R_c) that are bigger then 0. Indeed, the power-law is computed in log-log scale, therefore a mask on the values of $n(R, t') > 0$ is applied.

A.3.2 Transition Related Measures

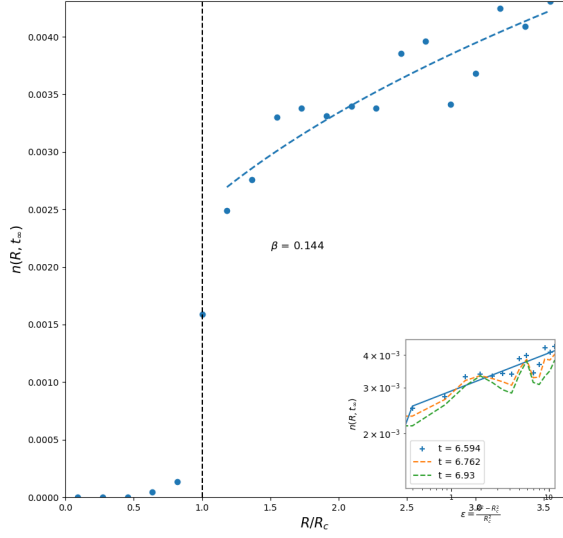
In this section we show how for different values of UCI in the case of Rio de Janeiro and San Francisco the observed transitions according to the methods and criteria introduced in [3.3.1](#). These plots highlights the dependence of such quantities on the monocentrism levels showing the dependence on the shape of the city for such quantities.



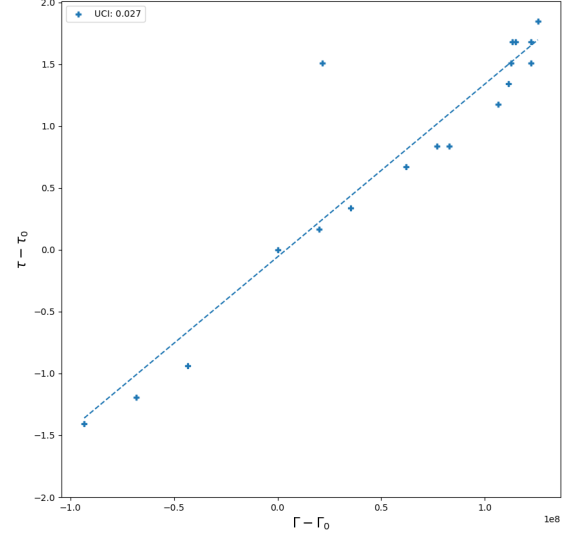
(a)



(b)

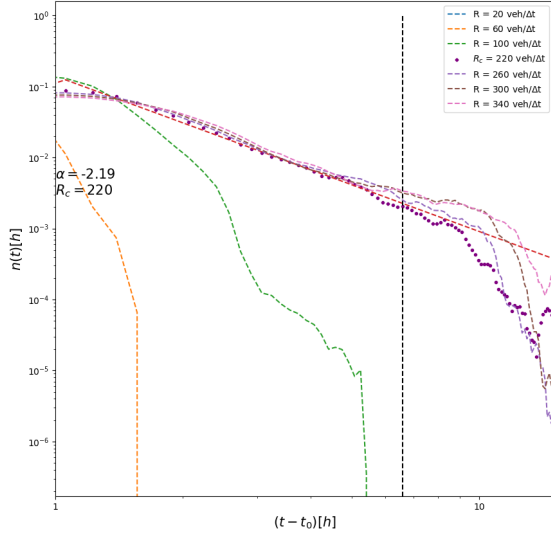


(c)

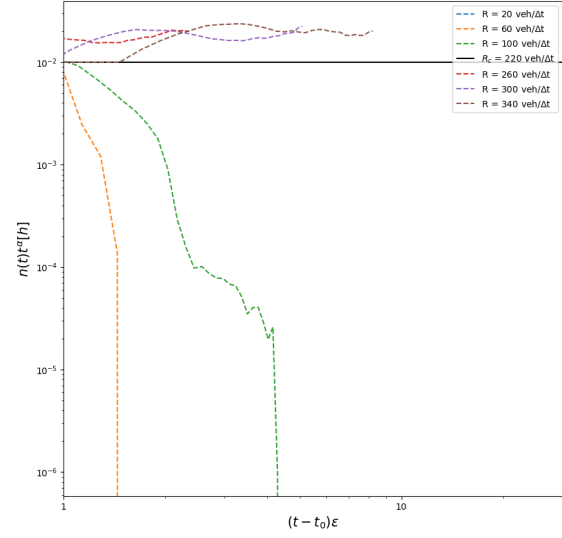


(d)

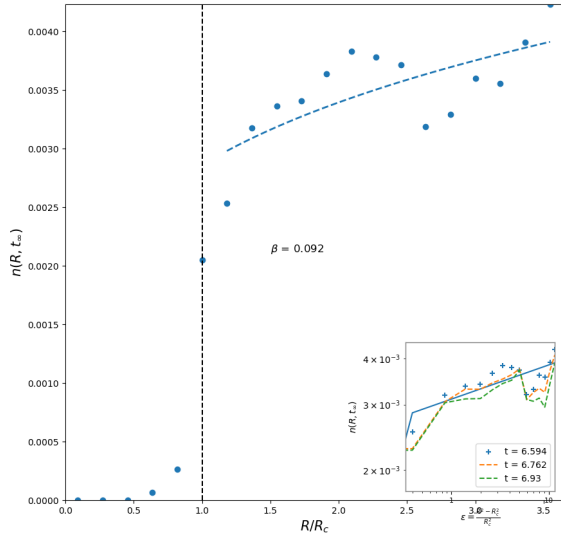
Figure A.13: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.027$. In A.13a is represented the number of people trapped in the road network at time t . In A.13b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.13c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.144$. In A.13d we see the linear increase of τ compared to Γ



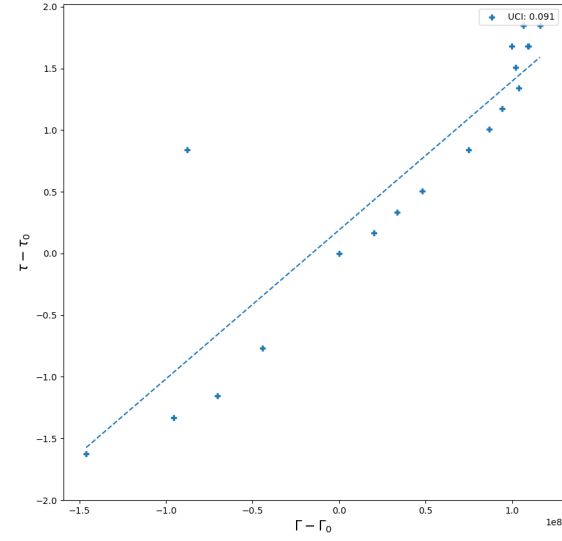
(a)



(b)

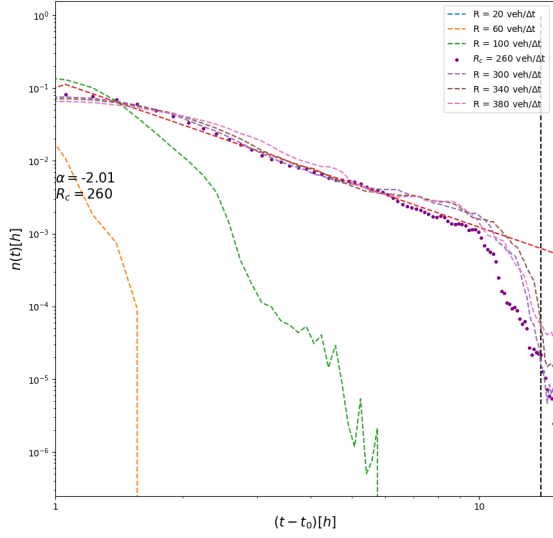


(c)

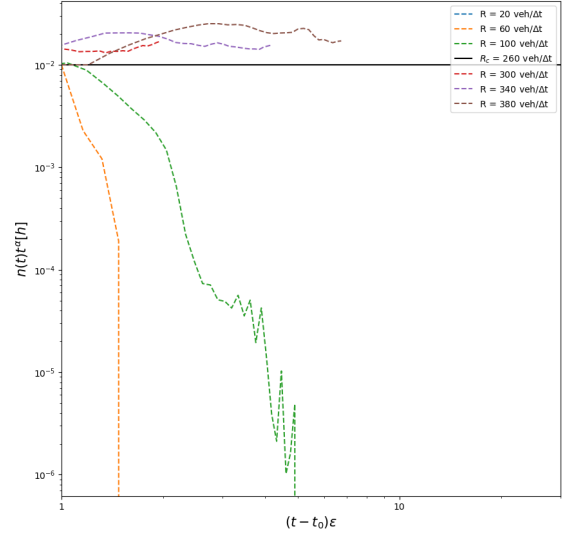


(d)

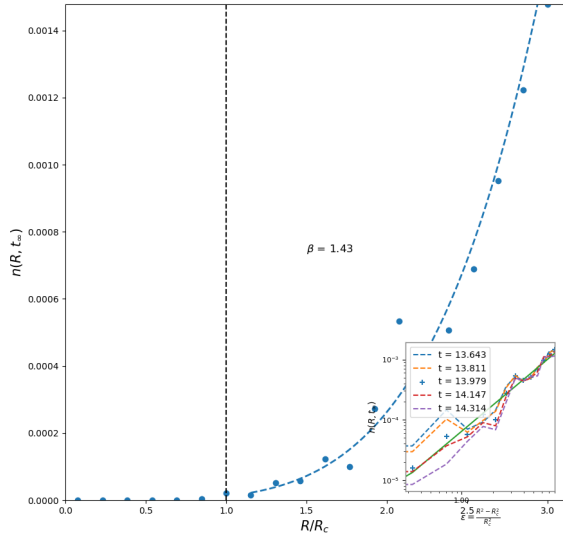
Figure A.14: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.091$. In A.14a is represented the number of people trapped in the road network at time t . In A.14b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.14c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.092$. In A.14d we see the linear increase of τ compared to Γ



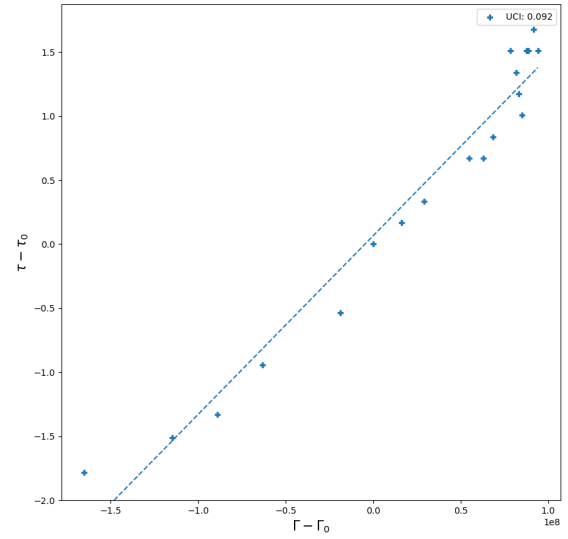
(a)



(b)

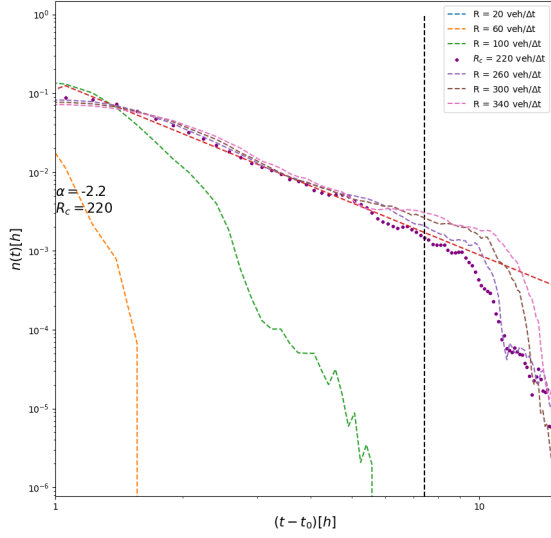


(c)

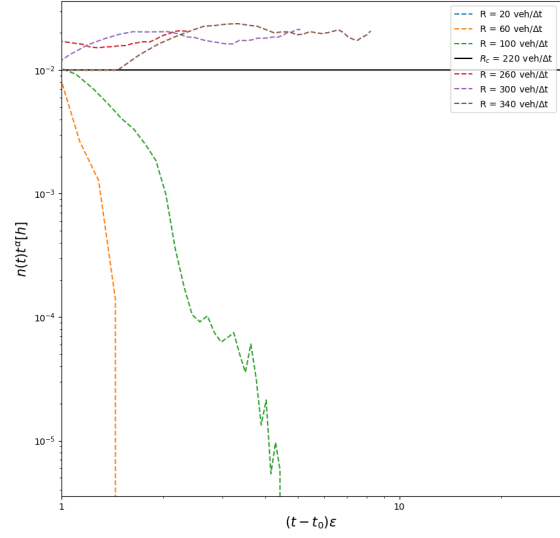


(d)

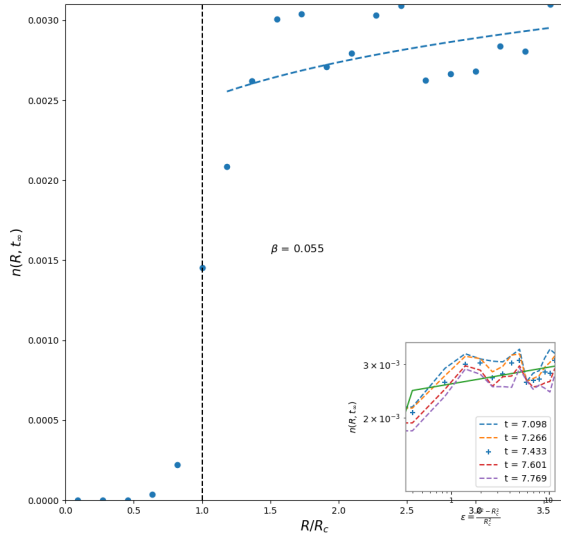
Figure A.15: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.092$. In A.15a is represented the number of people trapped in the road network at time t . In A.15b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.15c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.430$. In A.15d we see the linear increase of τ compared to Γ



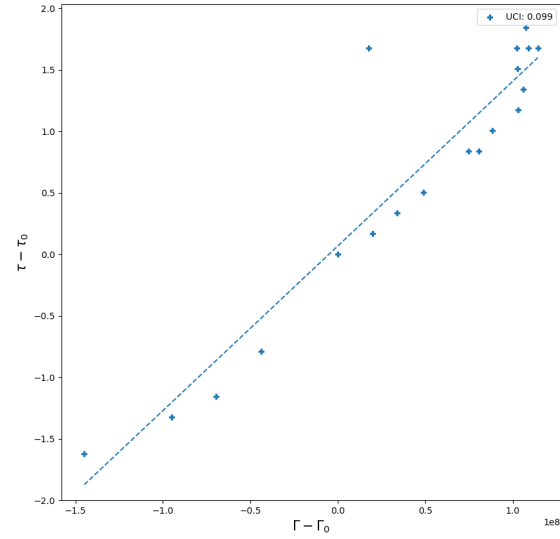
(a)



(b)

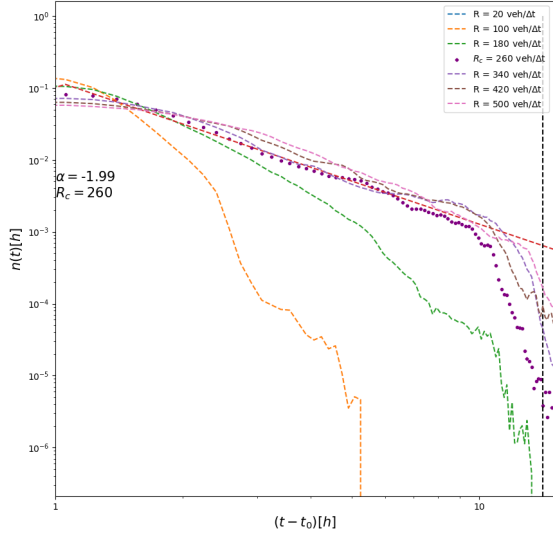


(c)

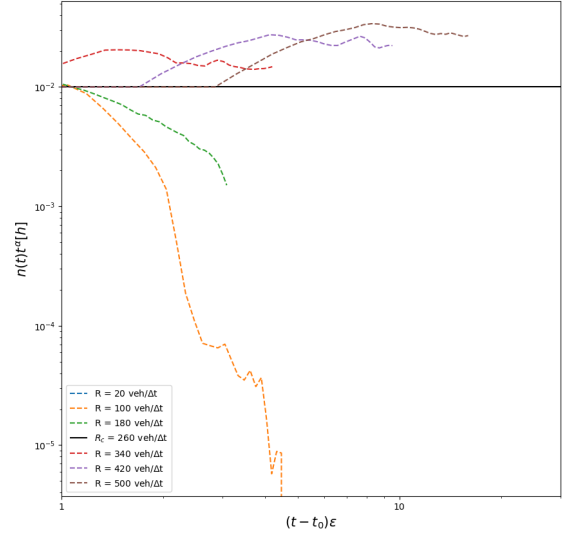


(d)

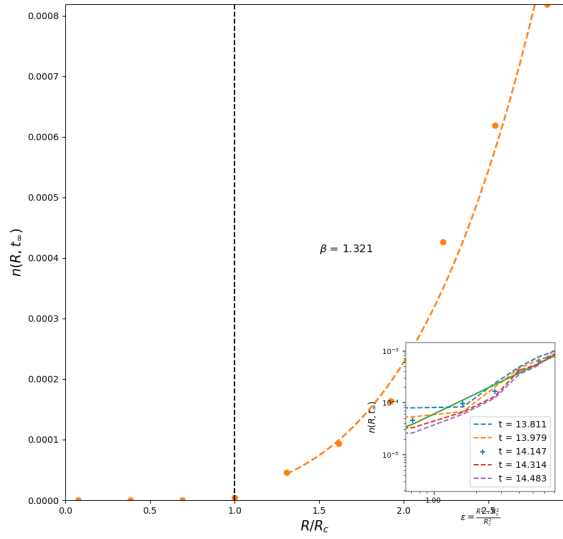
Figure A.16: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.099$. In A.16a is represented the number of people trapped in the road network at time t . In A.16b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.16c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.055$. In A.16d we see the linear increase of τ compared to Γ



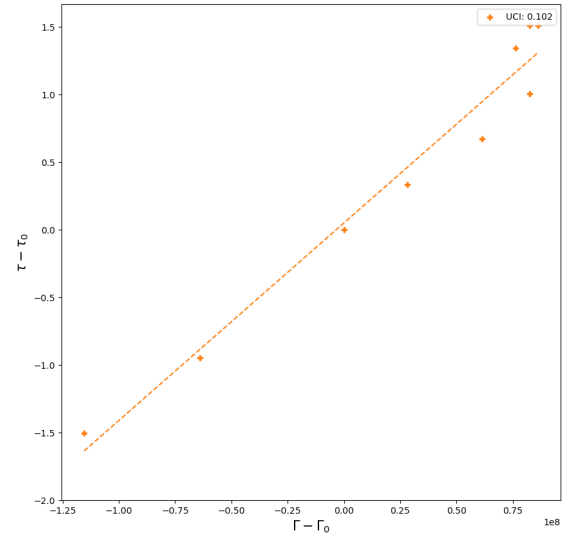
(a)



(b)

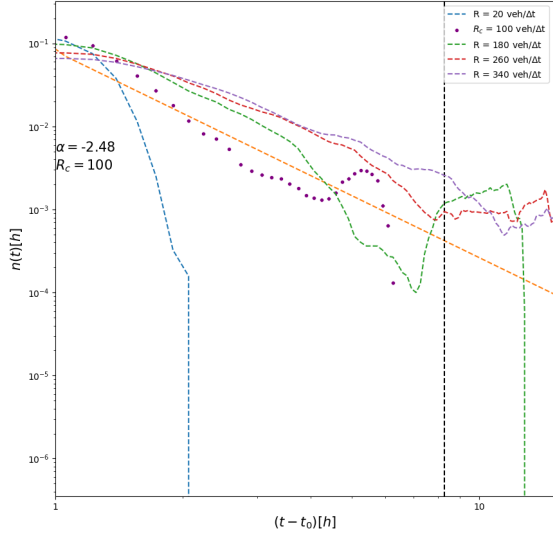


(c)

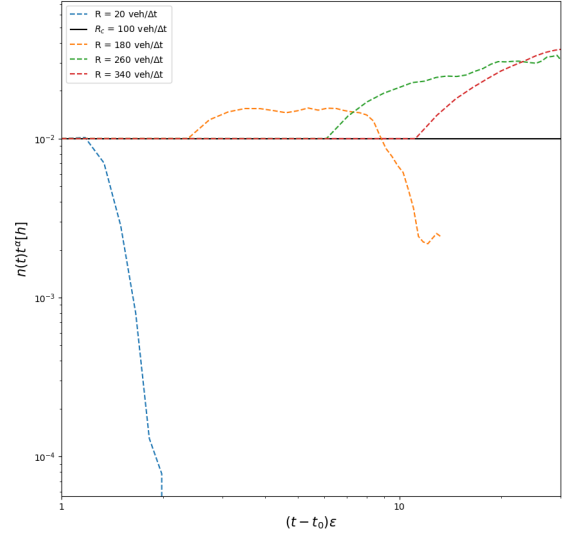


(d)

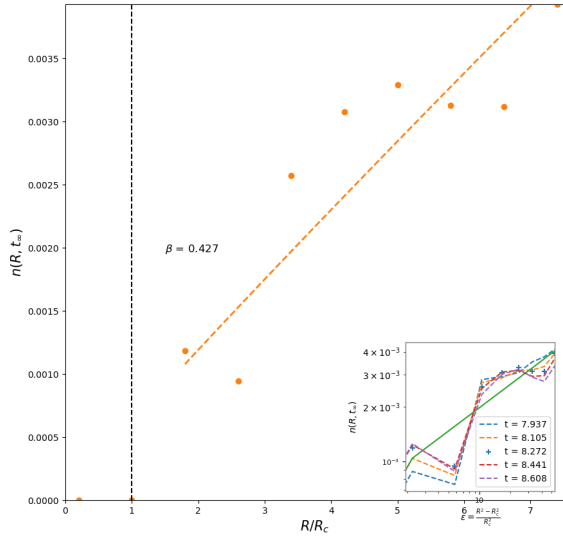
Figure A.17: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.102$. In A.17a is represented the number of people trapped in the road network at time t . In A.17b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.17c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.321$. In A.17d we see the linear increase of τ compared to Γ



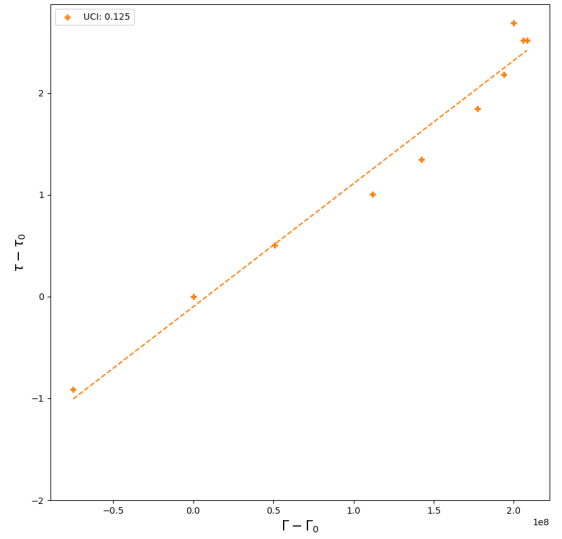
(a)



(b)

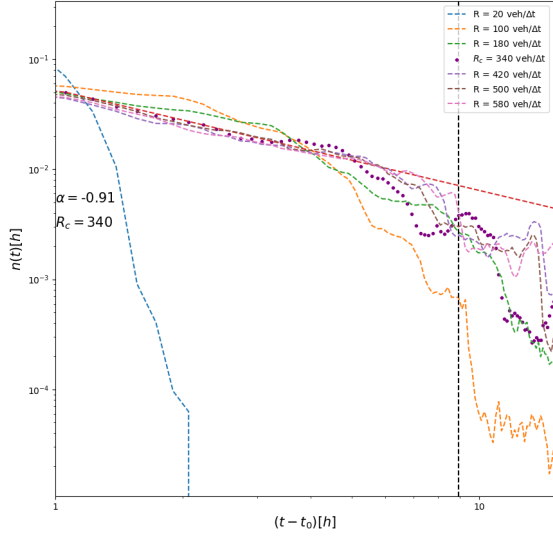


(c)

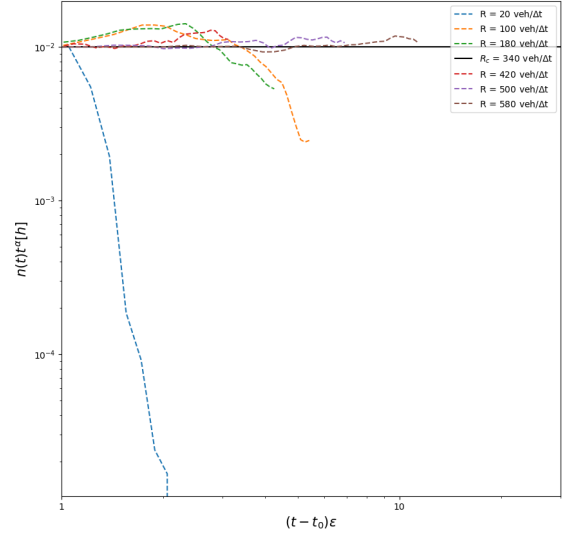


(d)

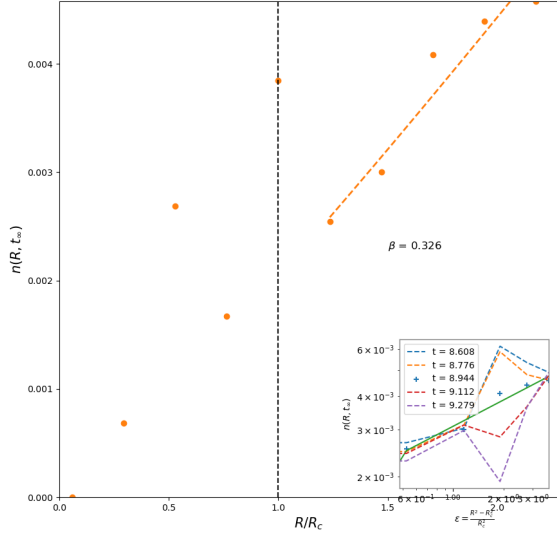
Figure A.18: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.125$. In A.18a is represented the number of people trapped in the road network at time t . In A.18b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.18c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.427$. In A.18d we see the linear increase of τ compared to Γ



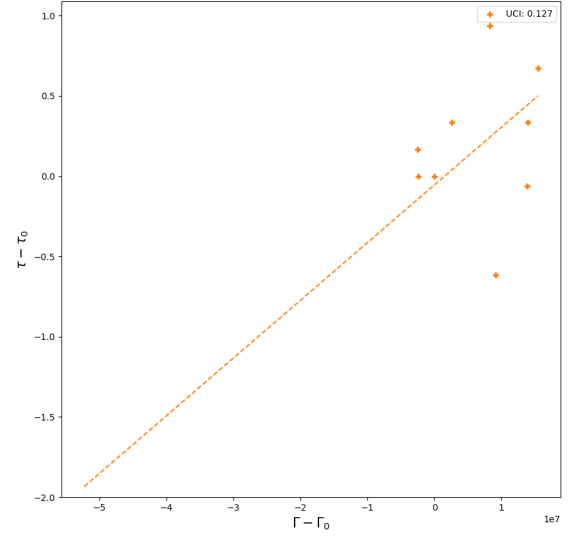
(a)



(b)

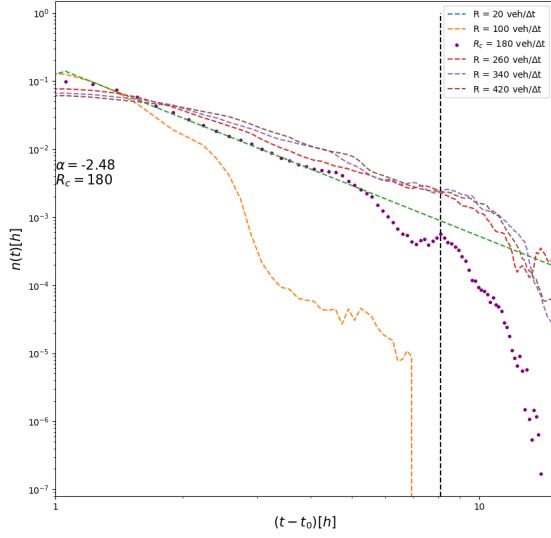


(c)

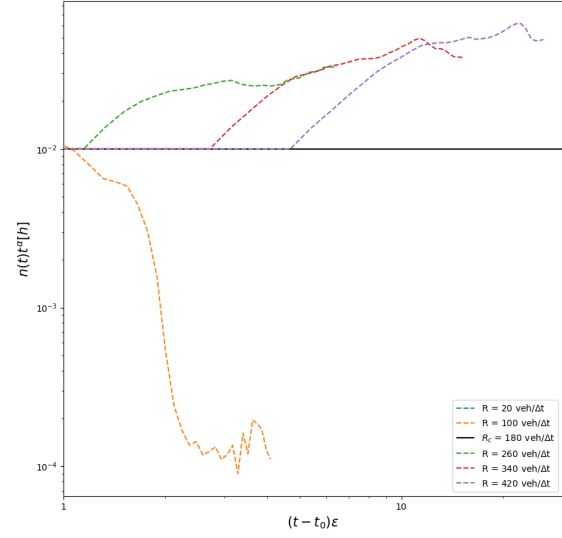


(d)

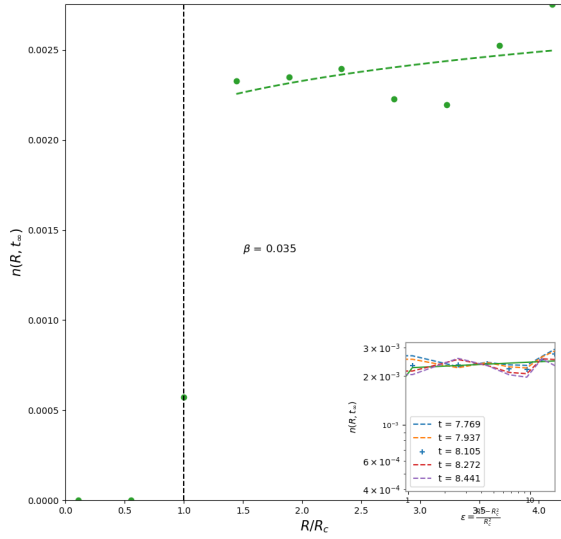
Figure A.19: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.127$. In A.19a is represented the number of people trapped in the road network at time t . In A.19b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.19c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.326$. In A.19d we see the linear increase of τ compared to Γ



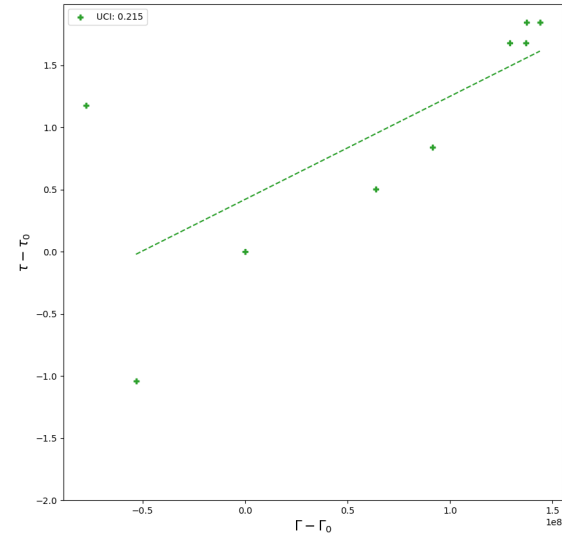
(a)



(b)

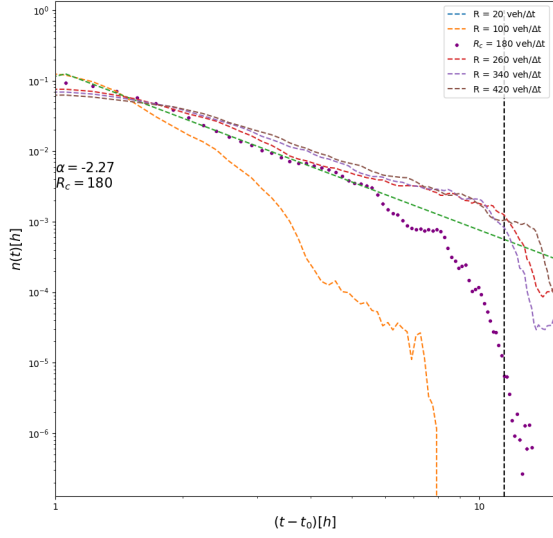


(c)

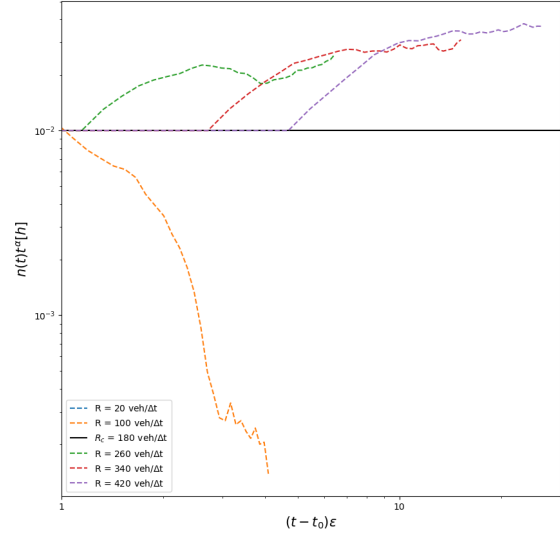


(d)

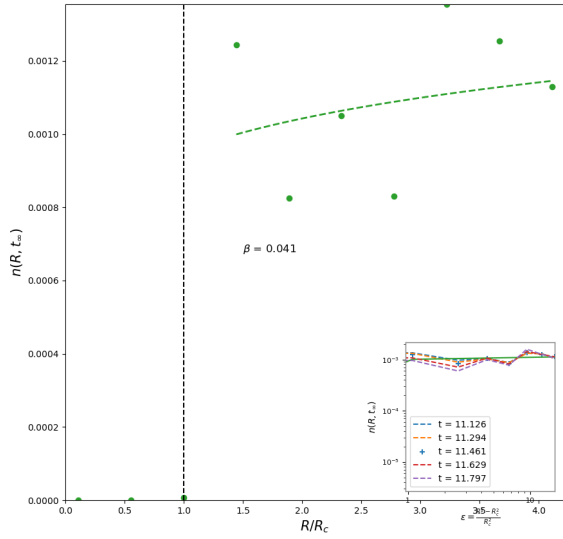
Figure A.20: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.215$. In A.20a is represented the number of people trapped in the road network at time t . In A.20b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.20c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.035$. In A.20d we see the linear increase of τ compared to Γ



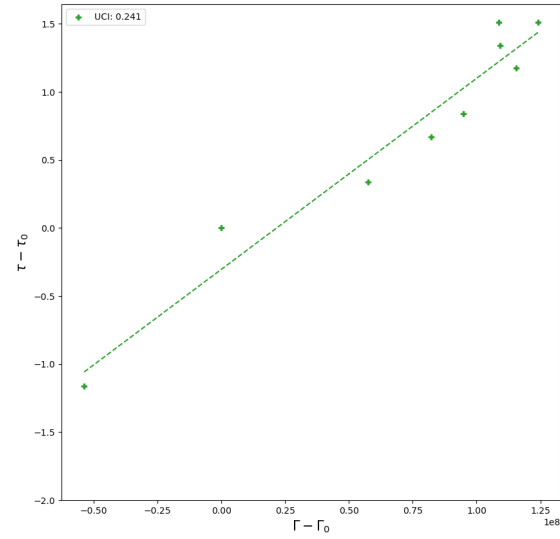
(a)



(b)

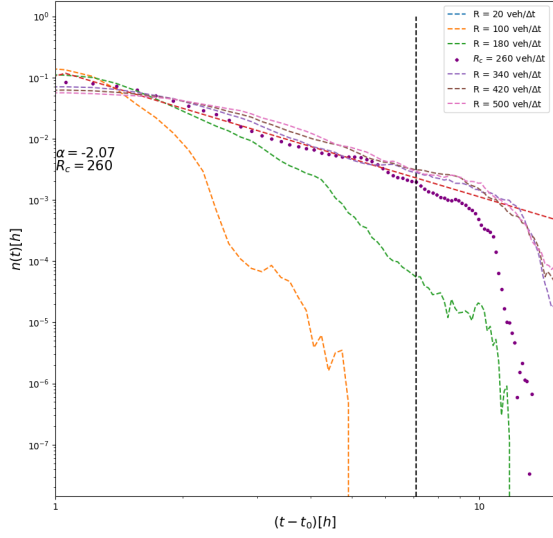


(c)

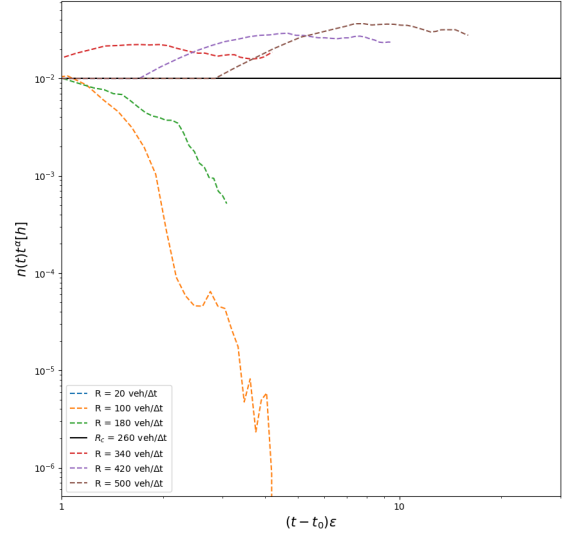


(d)

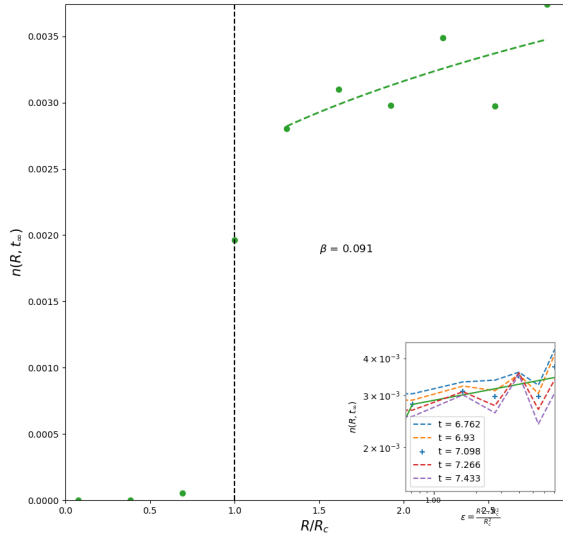
Figure A.21: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.241$. In A.21a is represented the number of people trapped in the road network at time t . In A.21b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.21c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.041$. In A.21d we see the linear increase of τ compared to Γ



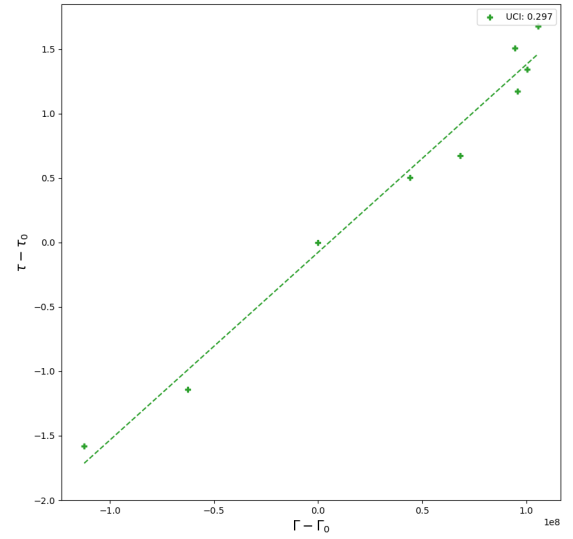
(a)



(b)

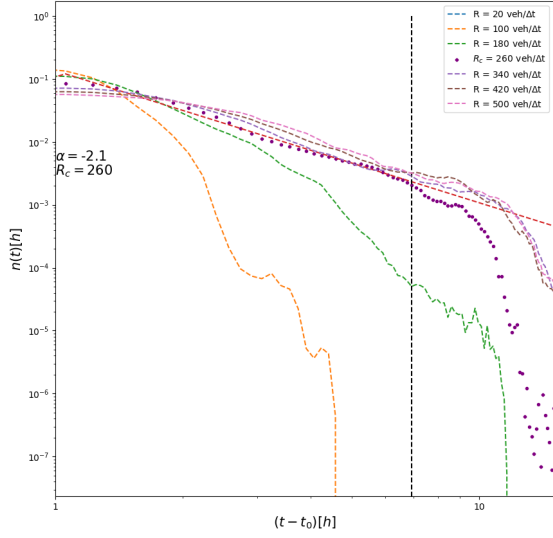


(c)

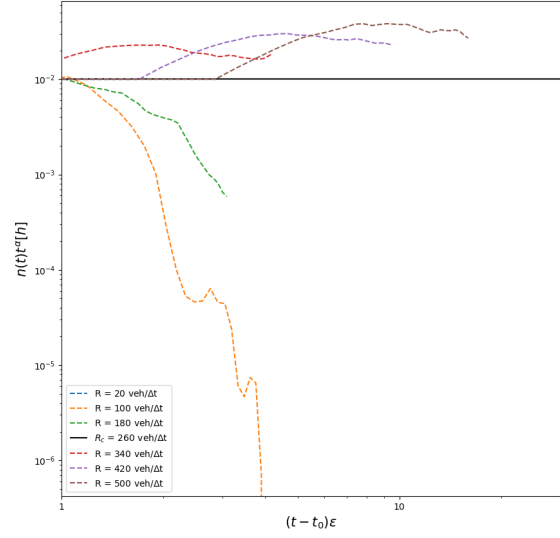


(d)

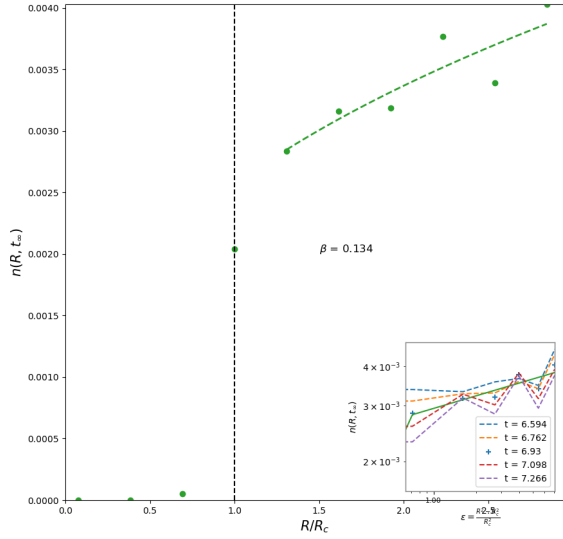
Figure A.22: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.297$. In A.22a is represented the number of people trapped in the road network at time t . In A.22b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.22c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.091$. In A.22d we see the linear increase of τ compared to Γ



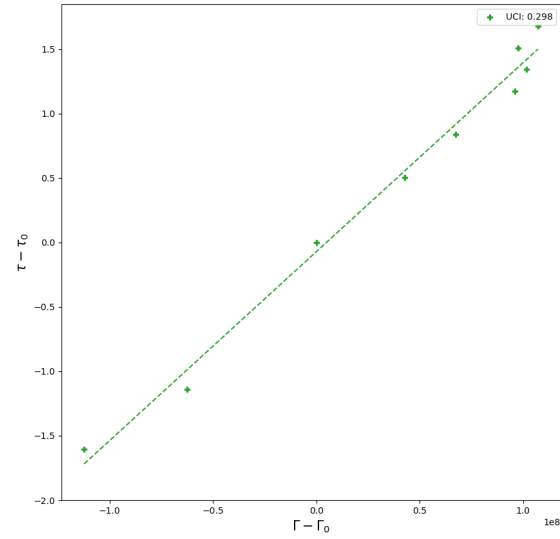
(a)



(b)

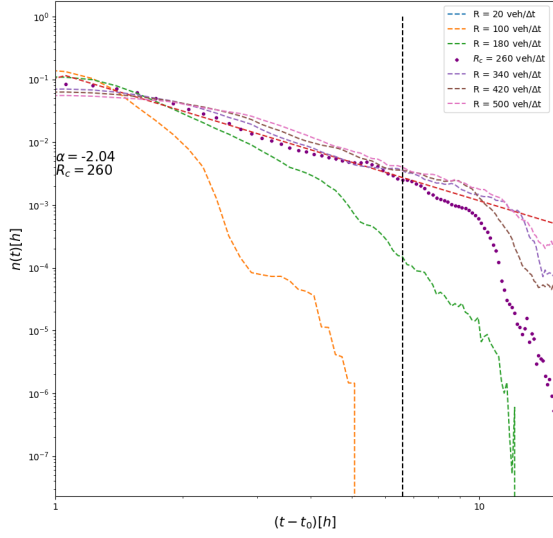


(c)

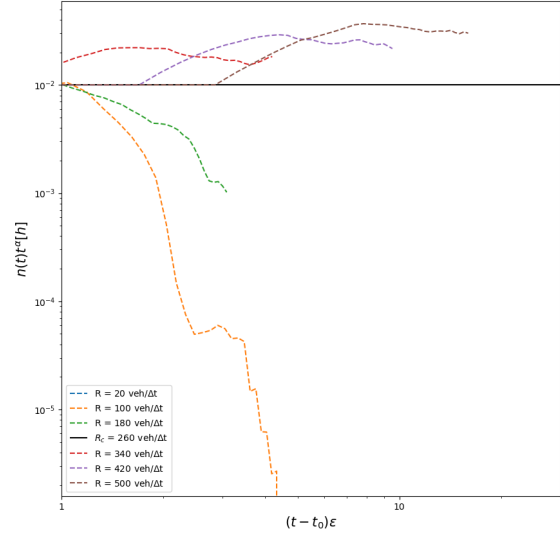


(d)

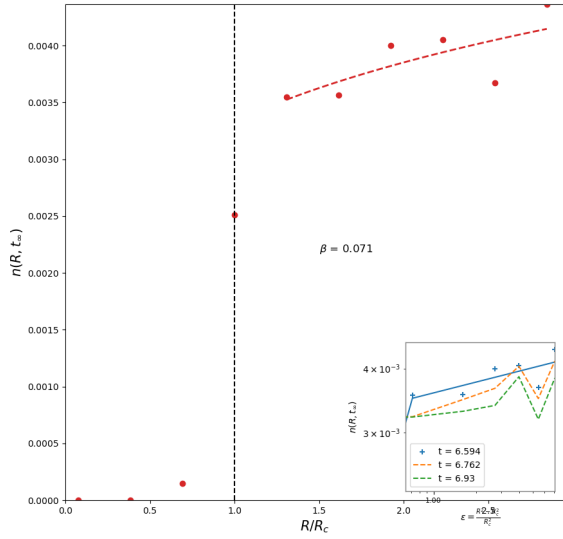
Figure A.23: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.298$. In A.23a is represented the number of people trapped in the road network at time t . In A.23b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.23c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.134$. In A.23d we see the linear increase of τ compared to Γ



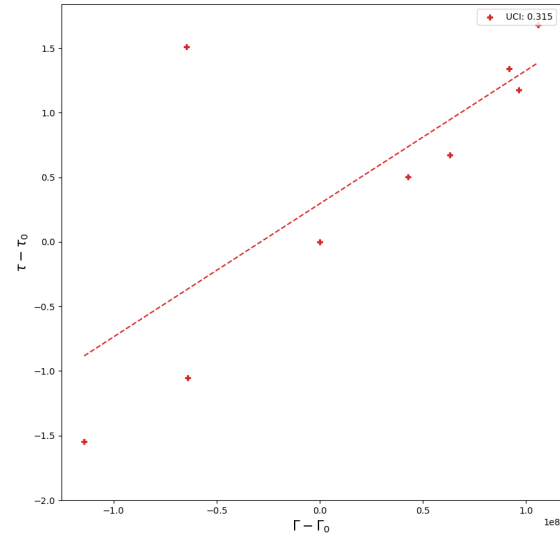
(a)



(b)

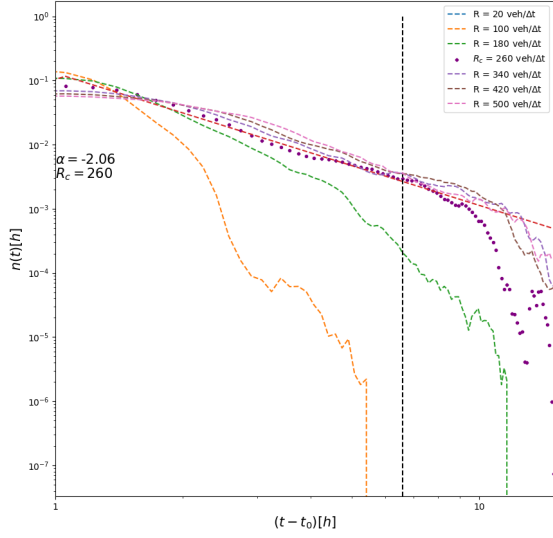


(c)

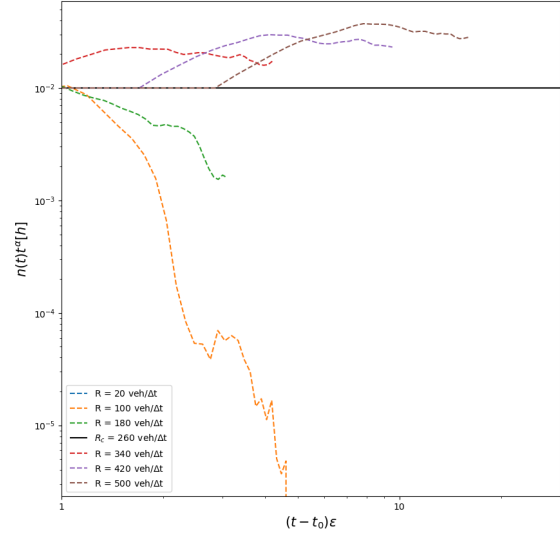


(d)

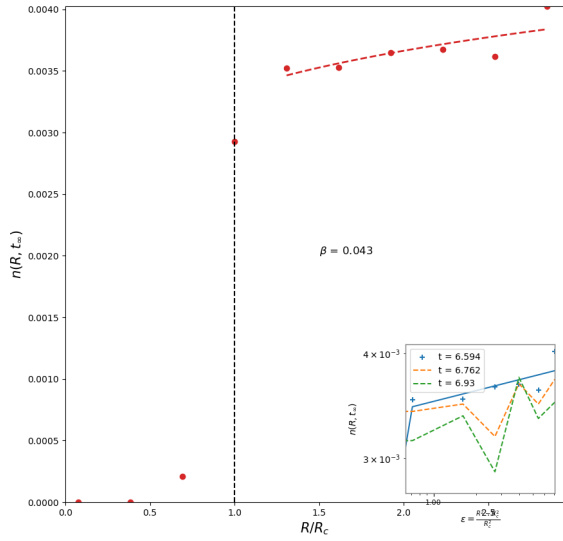
Figure A.24: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.315$. In A.24a is represented the number of people trapped in the road network at time t . In A.24b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.24c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.071$. In A.24d we see the linear increase of τ compared to Γ



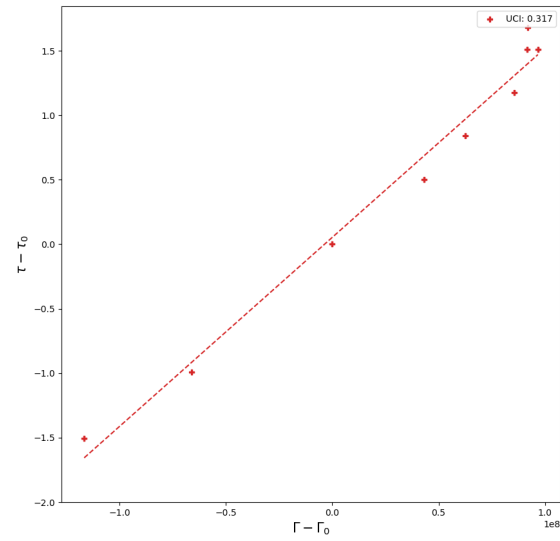
(a)



(b)

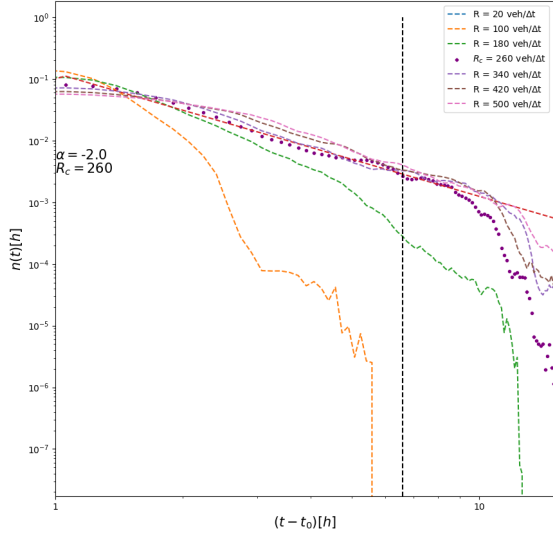


(c)

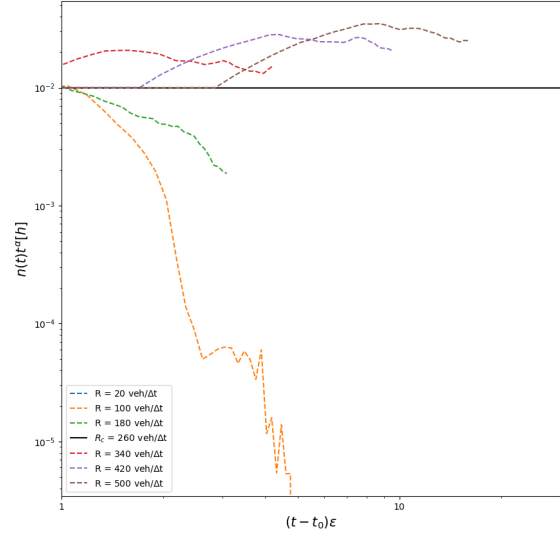


(d)

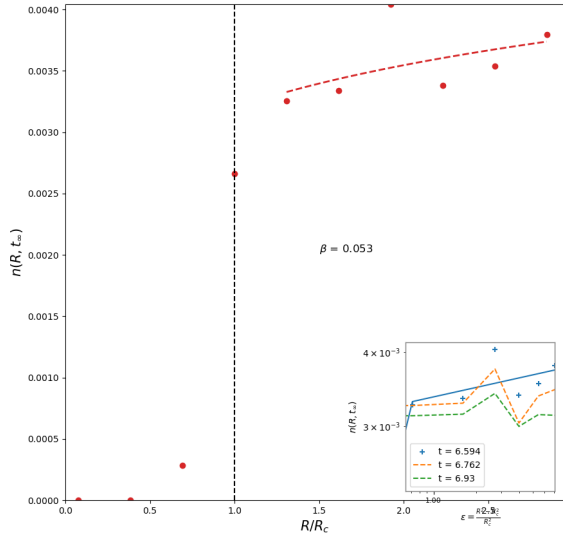
Figure A.25: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.317$. In A.25a is represented the number of people trapped in the road network at time t . In A.25b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.25c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.043$. In A.25d we see the linear increase of τ compared to Γ



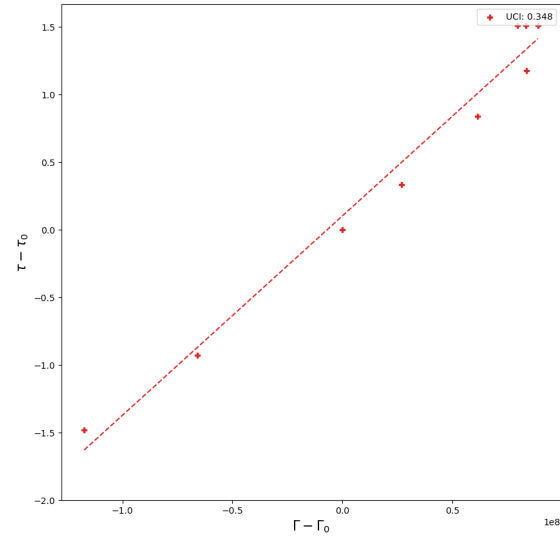
(a)



(b)

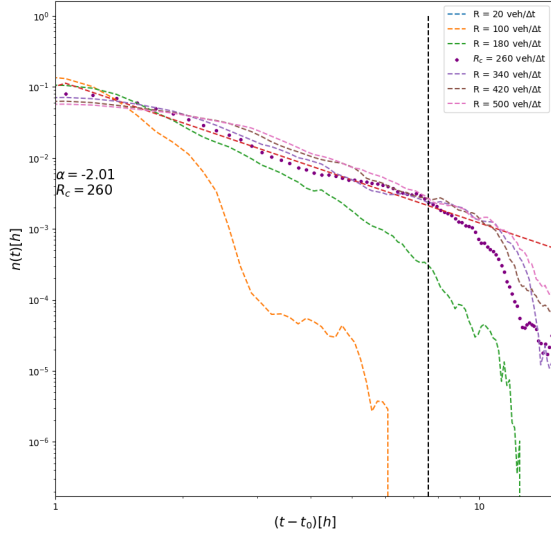


(c)

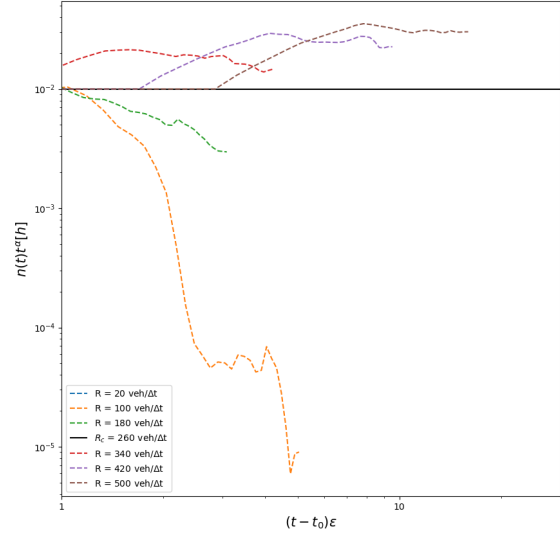


(d)

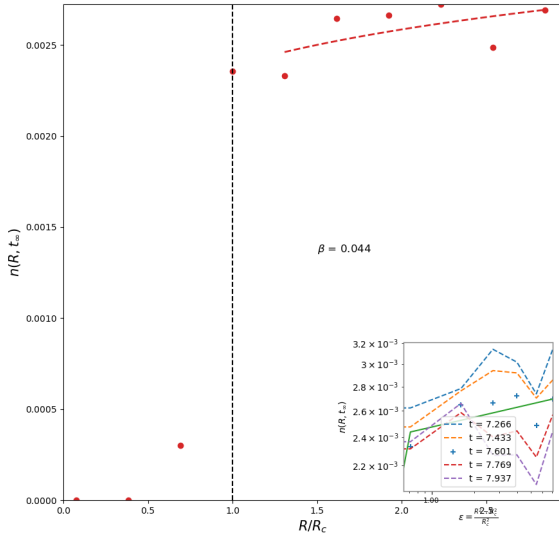
Figure A.26: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.348$. In A.26a is represented the number of people trapped in the road network at time t . In A.26b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.26c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.053$. In A.26d we see the linear increase of τ compared to Γ



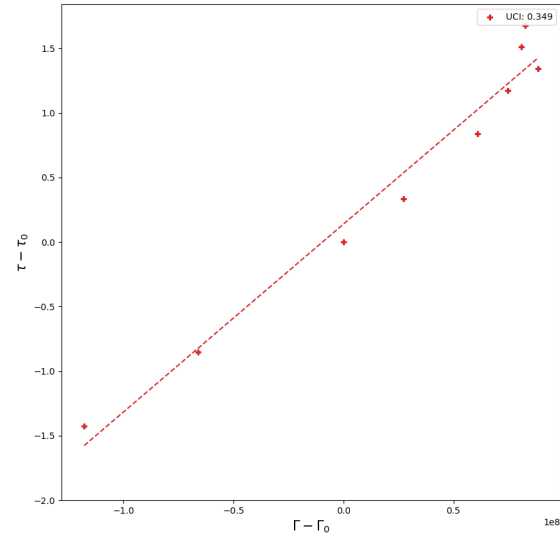
(a)



(b)

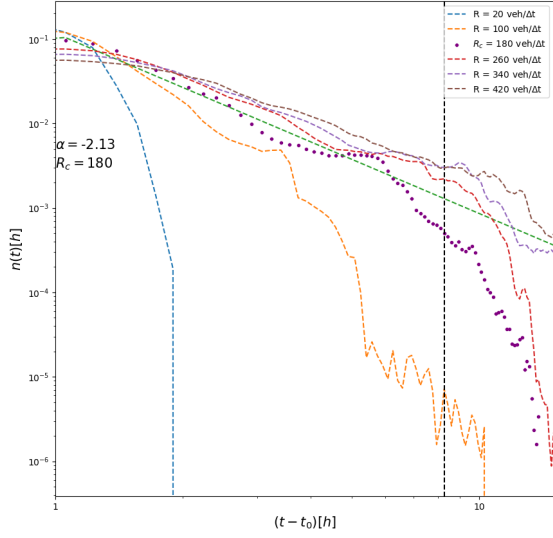


(c)

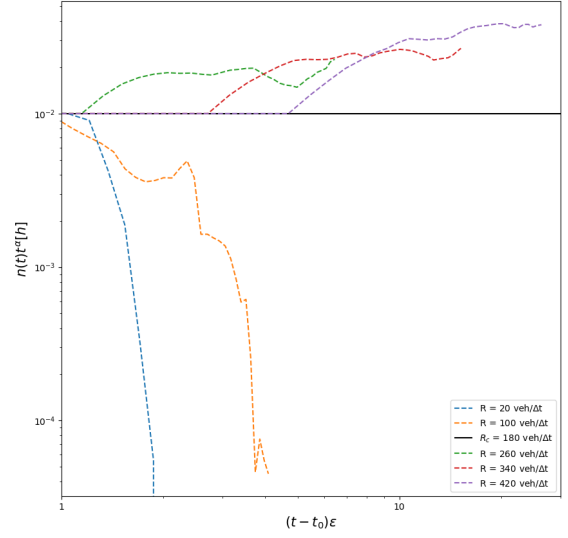


(d)

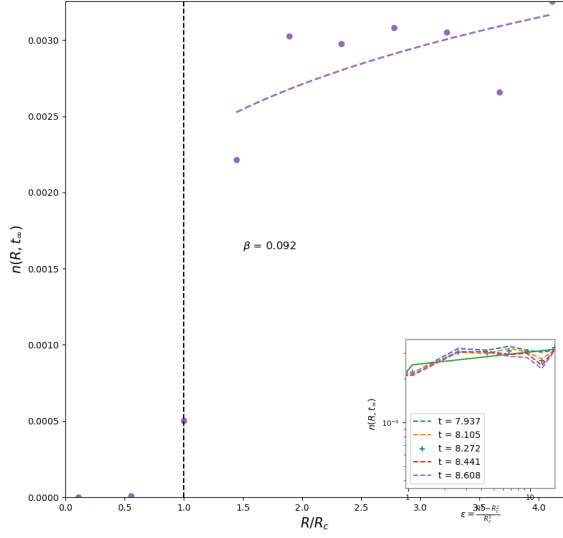
Figure A.27: Here are represented the important features about the phase transition in Rio De Janeiro for UCI = 0.349. In A.27a is represented the number of people trapped in the road network at time t . In A.27b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.27c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.044$. In A.27d we see the linear increase of τ compared to Γ



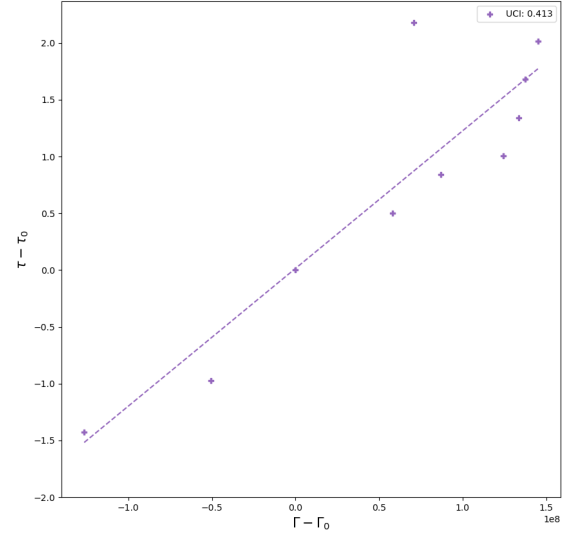
(a)



(b)

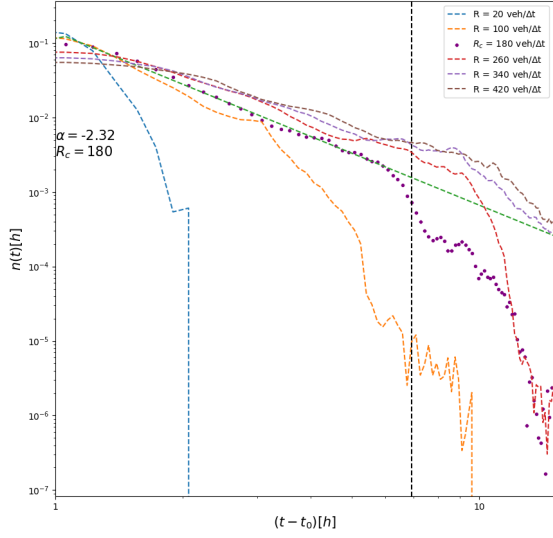


(c)

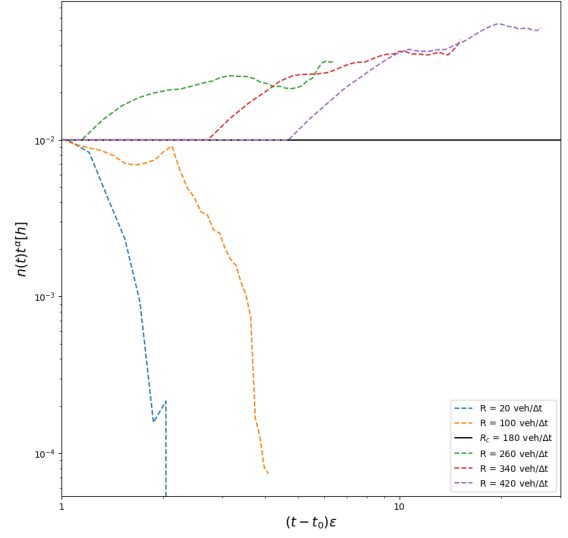


(d)

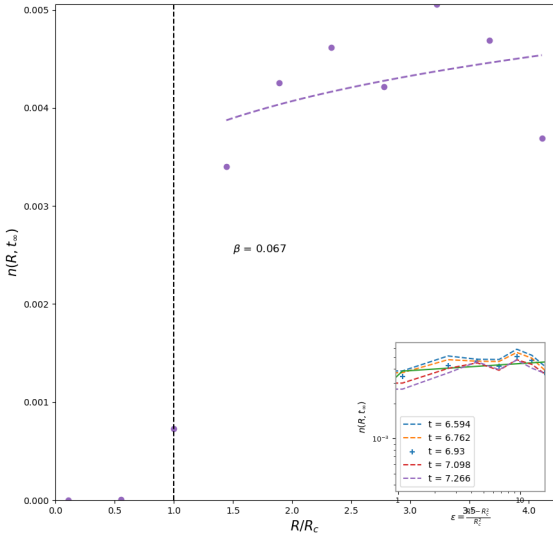
Figure A.28: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.413$. In A.28a is represented the number of people trapped in the road network at time t . In A.28b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.28c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.092$. In A.28d we see the linear increase of τ compared to Γ



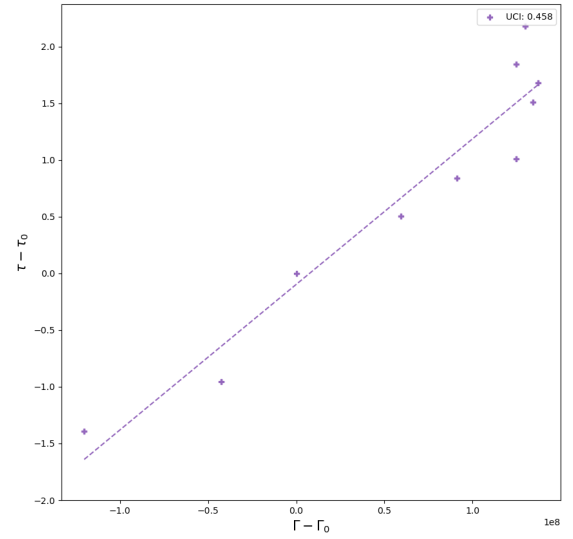
(a)



(b)

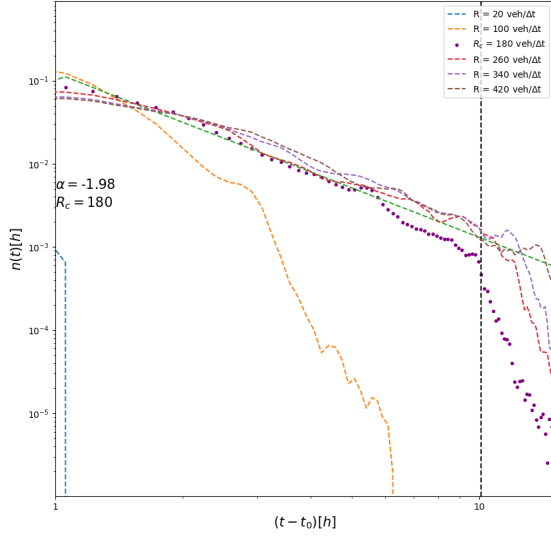


(c)

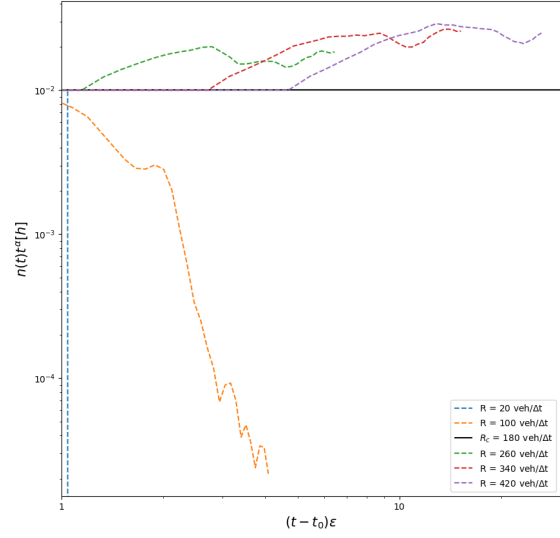


(d)

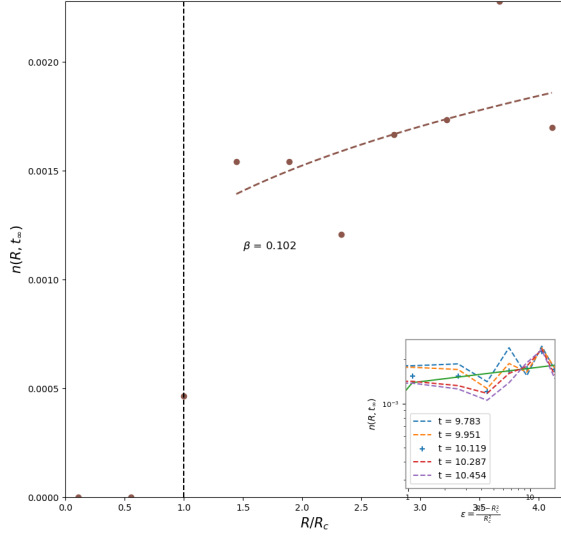
Figure A.29: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.458$. In A.29a is represented the number of people trapped in the road network at time t . In A.29b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.29c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.067$. In A.29d we see the linear increase of τ compared to Γ



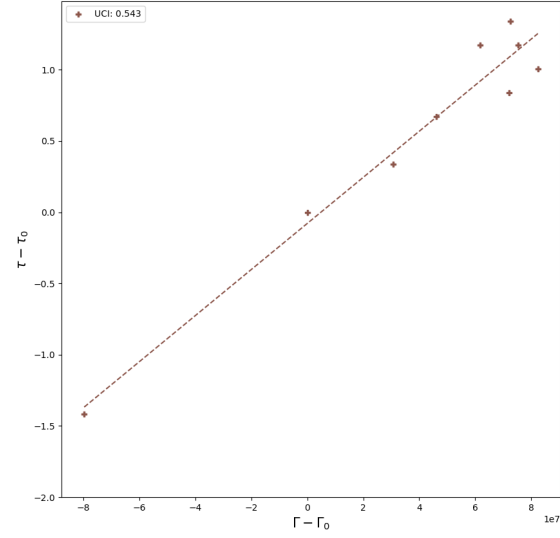
(a)



(b)

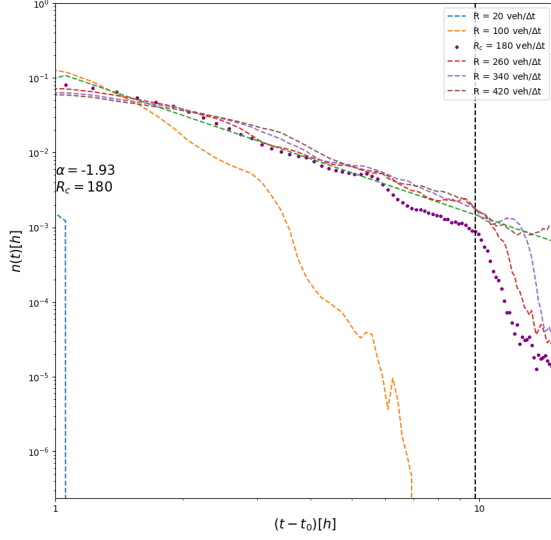


(c)

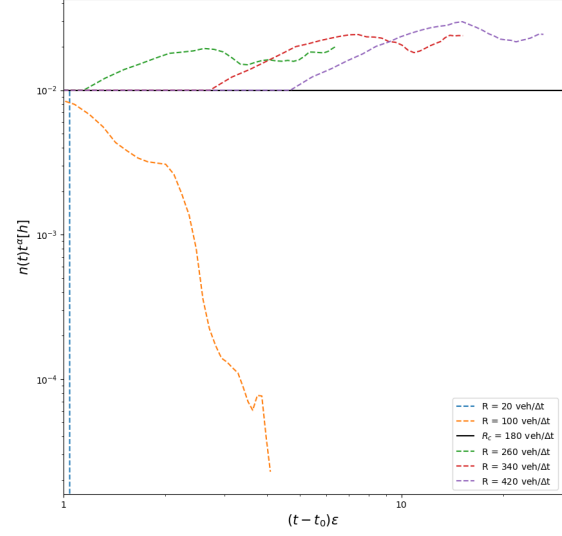


(d)

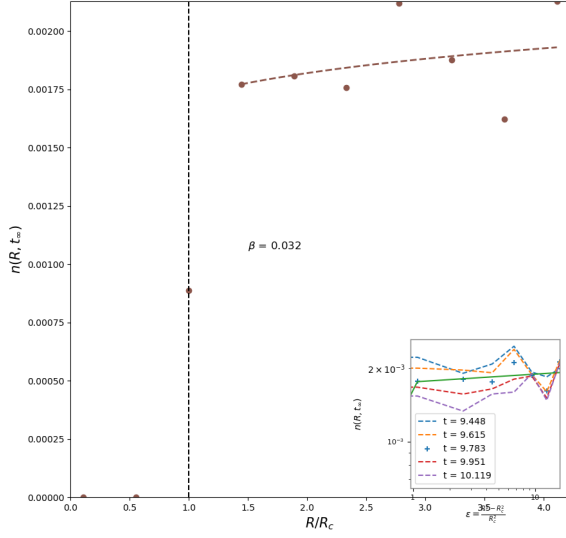
Figure A.30: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.543$. In A.30a is represented the number of people trapped in the road network at time t . In A.30b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.30c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.102$. In A.30d we see the linear increase of τ compared to Γ



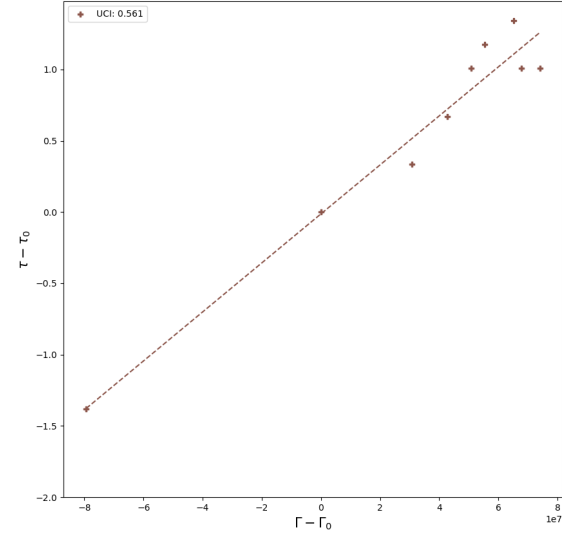
(a)



(b)

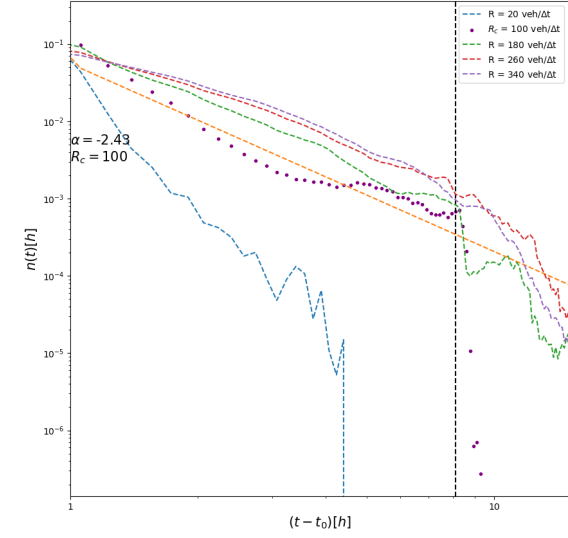


(c)

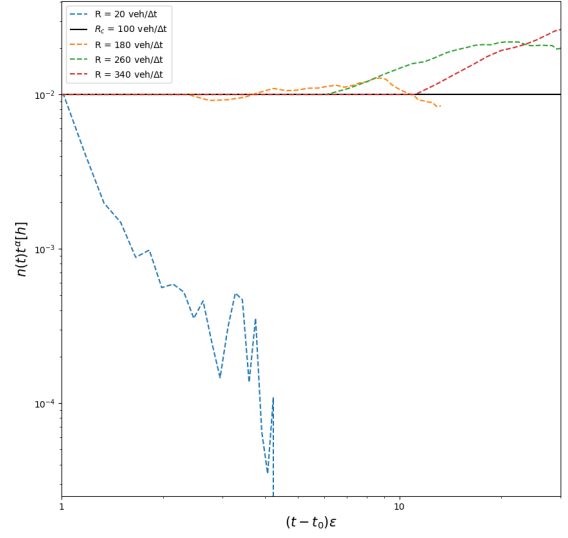


(d)

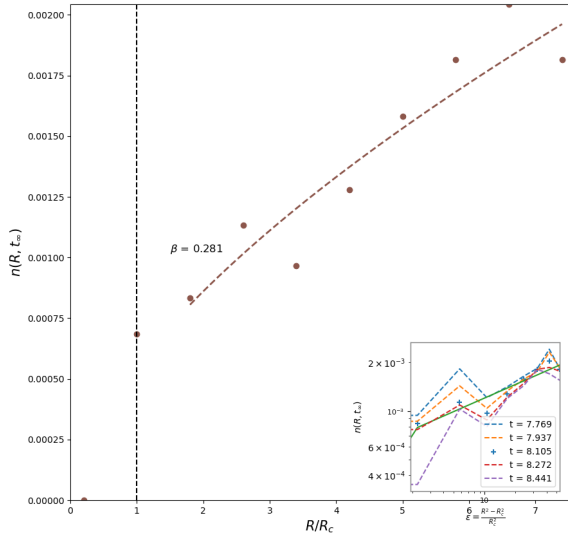
Figure A.31: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.561$. In A.31a is represented the number of people trapped in the road network at time t . In A.31b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.31c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.032$. In A.31d we see the linear increase of τ compared to Γ



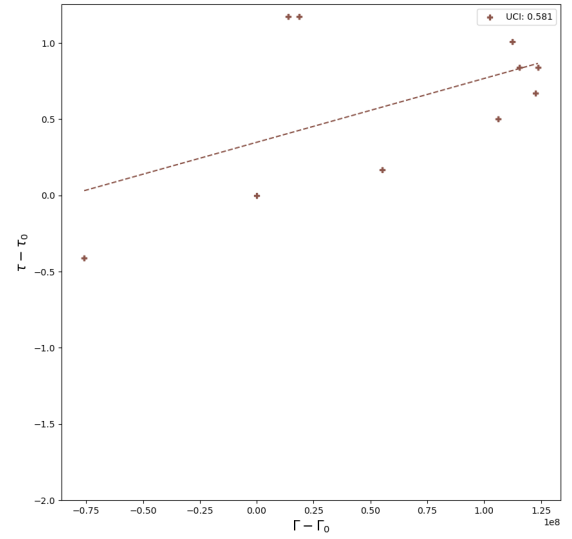
(a)



(b)

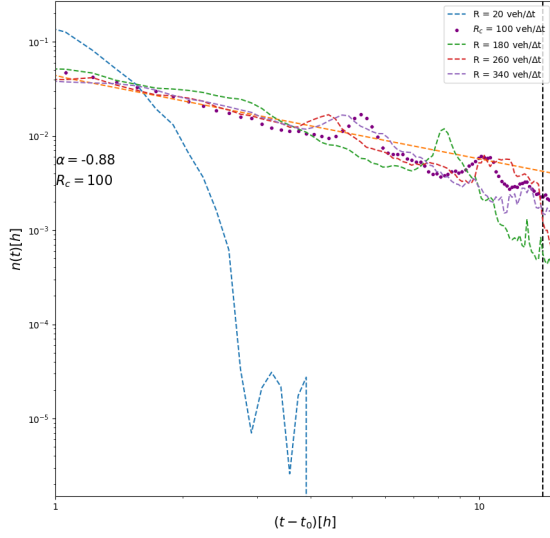


(c)

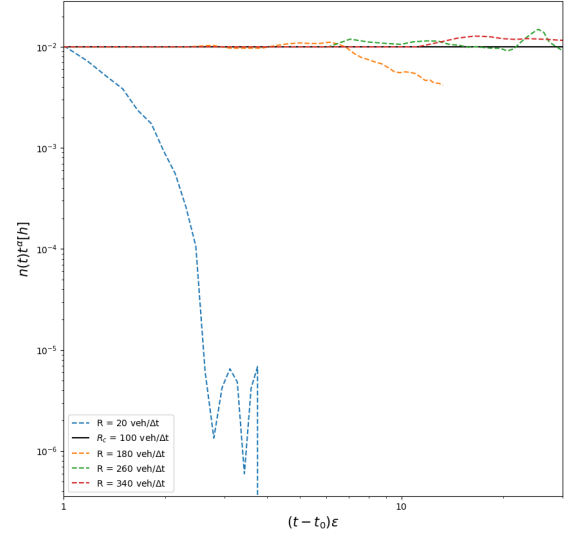


(d)

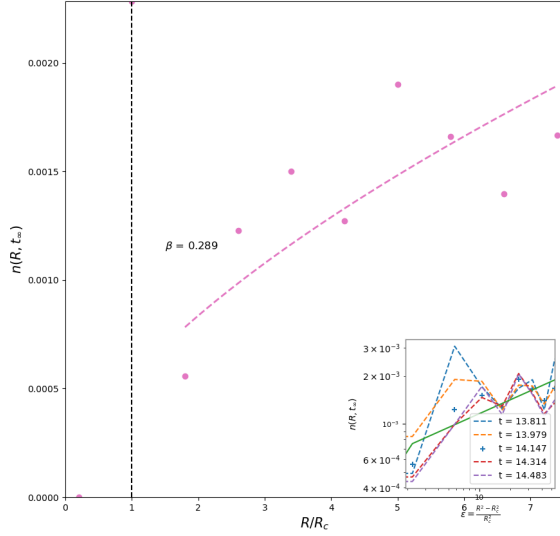
Figure A.32: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.581$. In A.32a is represented the number of people trapped in the road network at time t . In A.32b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.32c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.281$. In A.32d we see the linear increase of τ compared to Γ



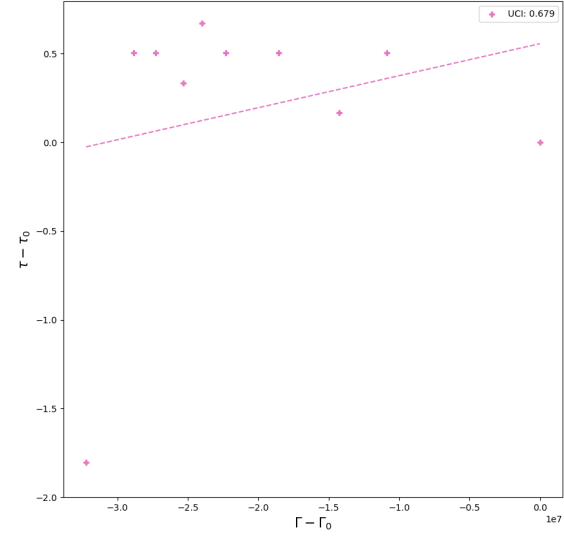
(a)



(b)

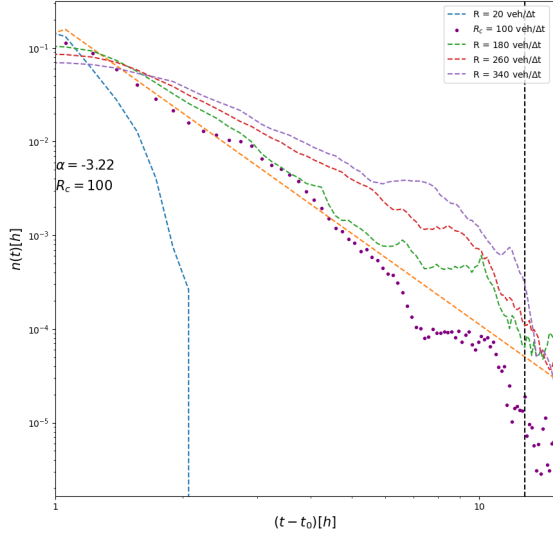


(c)

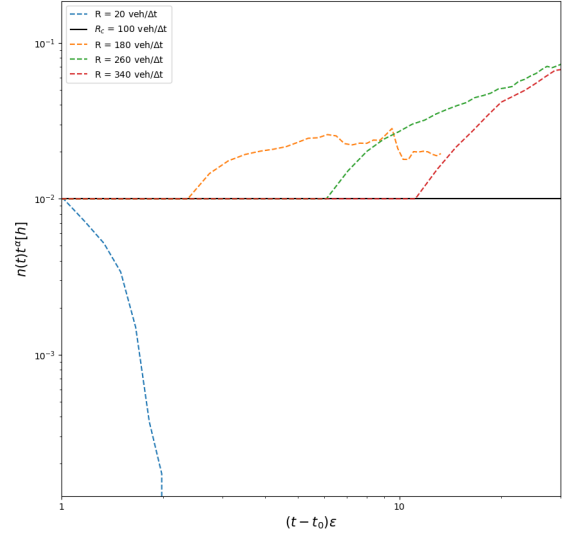


(d)

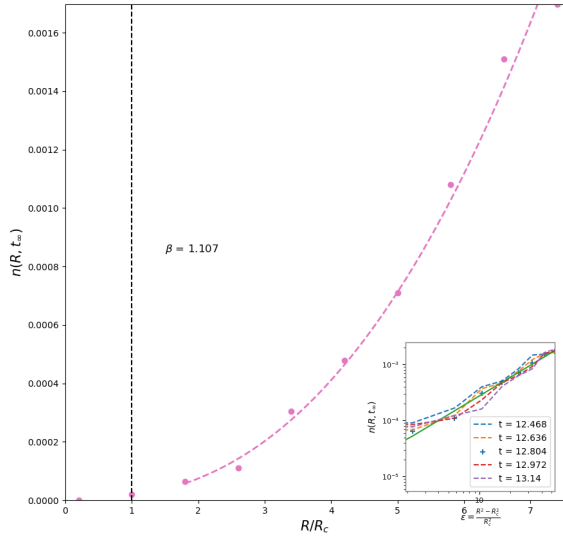
Figure A.33: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.679$. In A.33a is represented the number of people trapped in the road network at time t . In A.33b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.33c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.289$. In A.33d we see the linear increase of τ compared to Γ



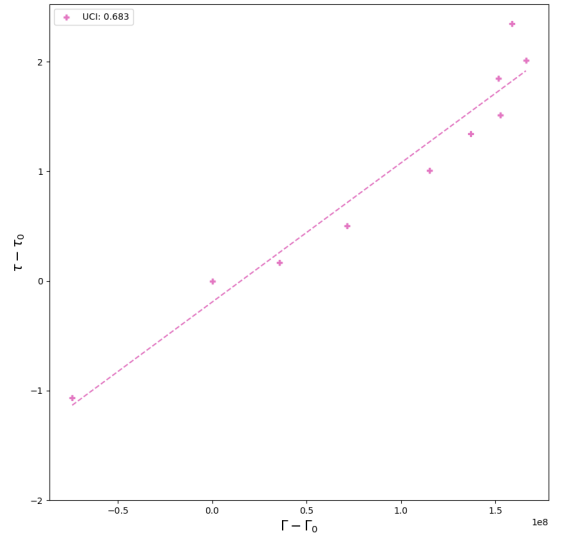
(a)



(b)



(c)



(d)

Figure A.34: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.683$. In A.34a is represented the number of people trapped in the road network at time t . In A.34b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.34c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.107$. In A.34d we see the linear increase of τ compared to Γ

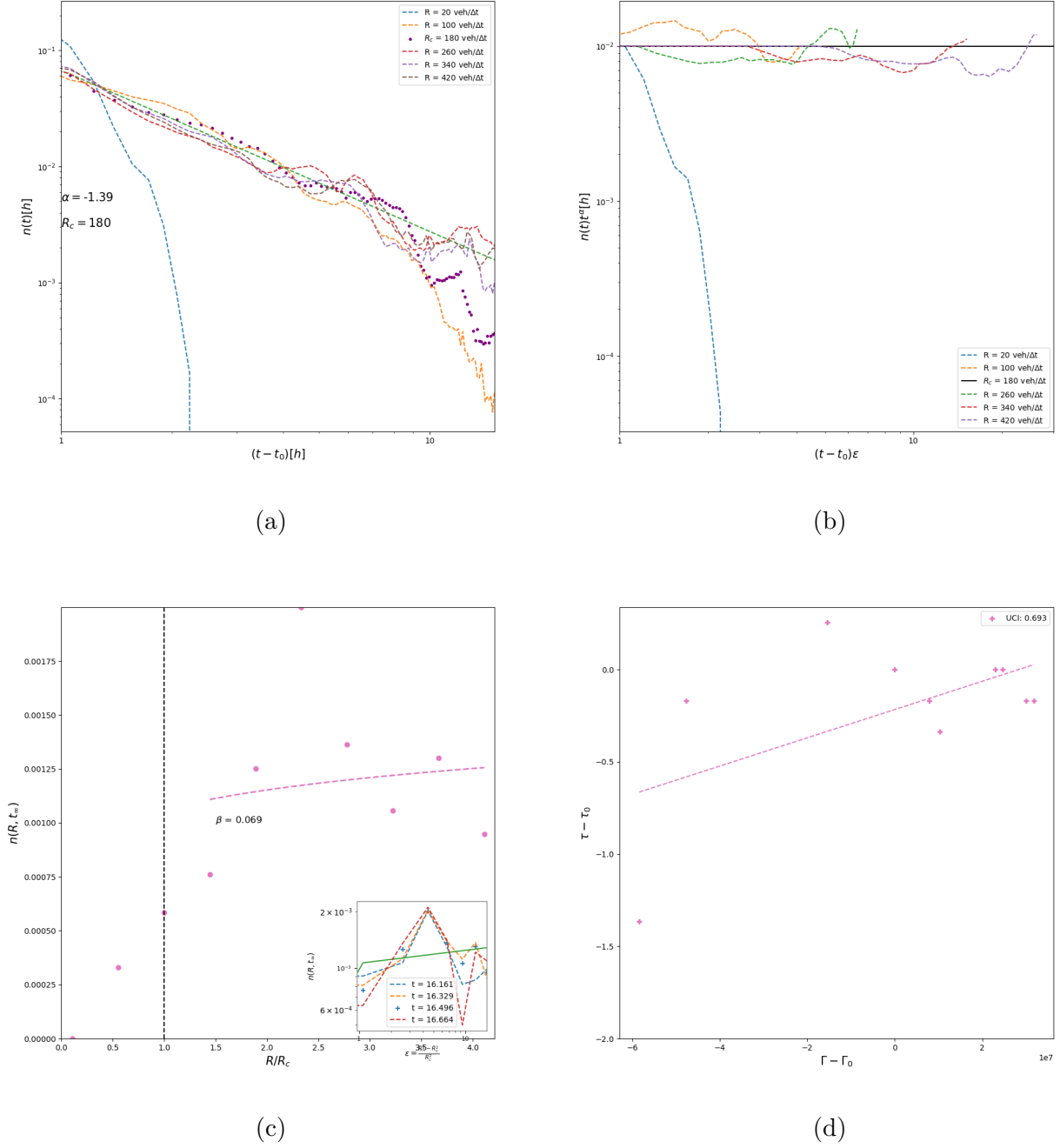
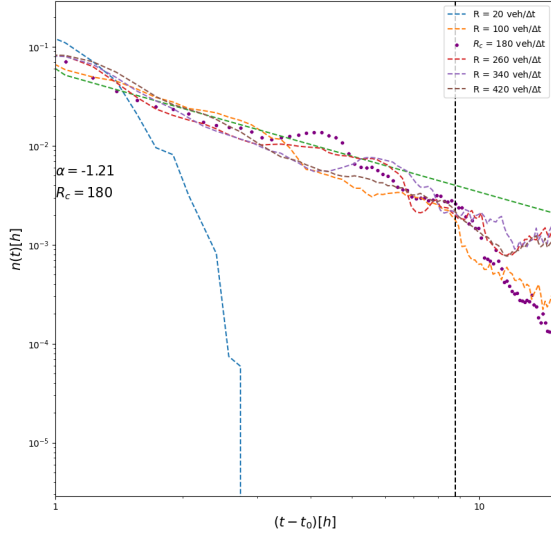
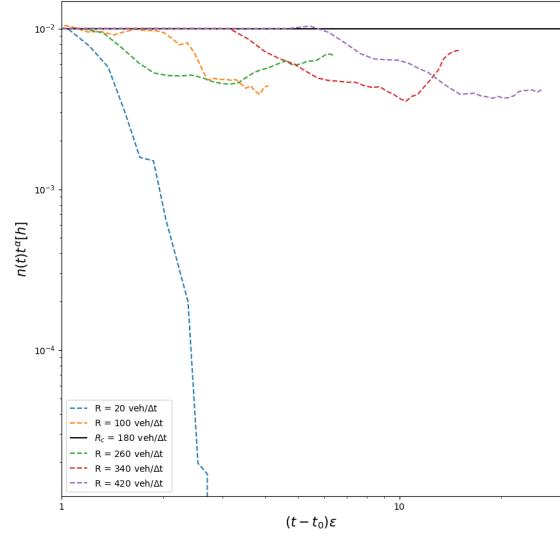


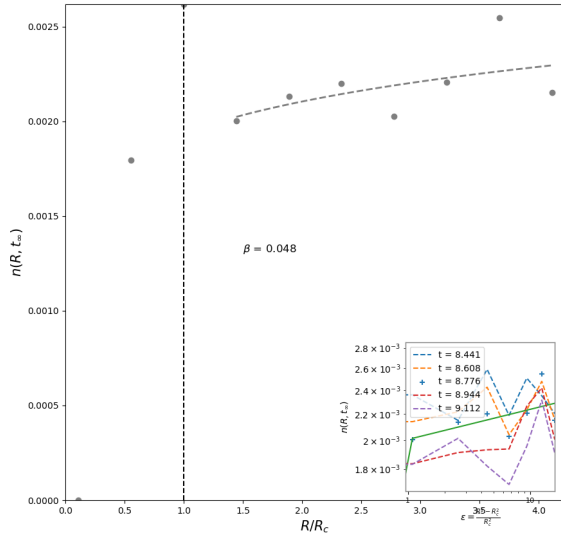
Figure A.35: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.693$. In A.35a is represented the number of people trapped in the road network at time t . In A.35b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.35c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.069$. In A.35d we see the linear increase of τ compared to Γ



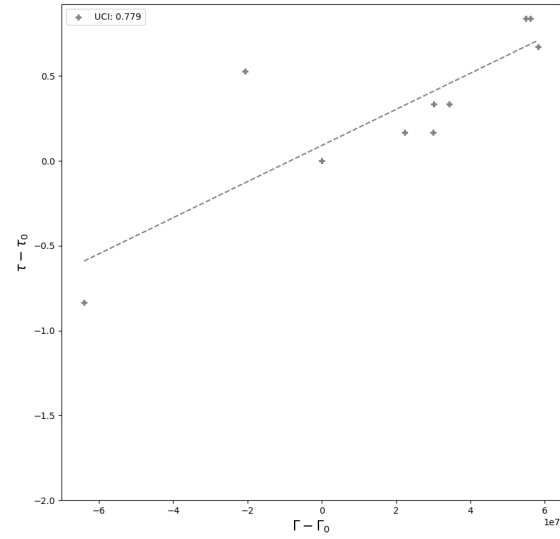
(a)



(b)

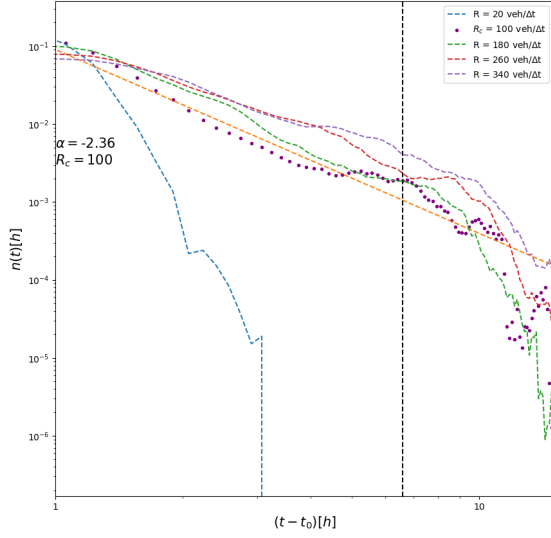


(c)

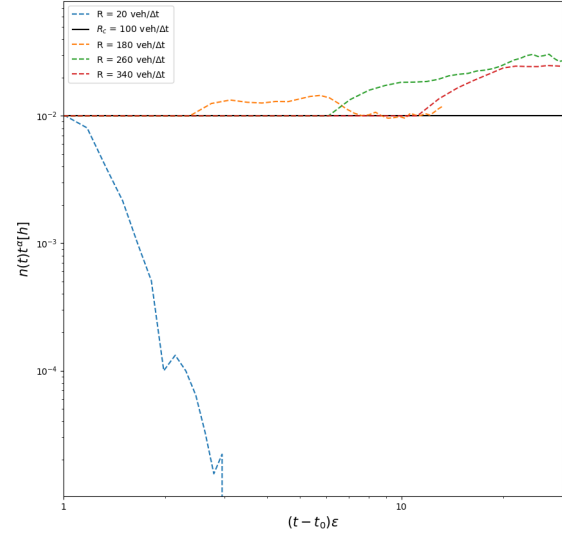


(d)

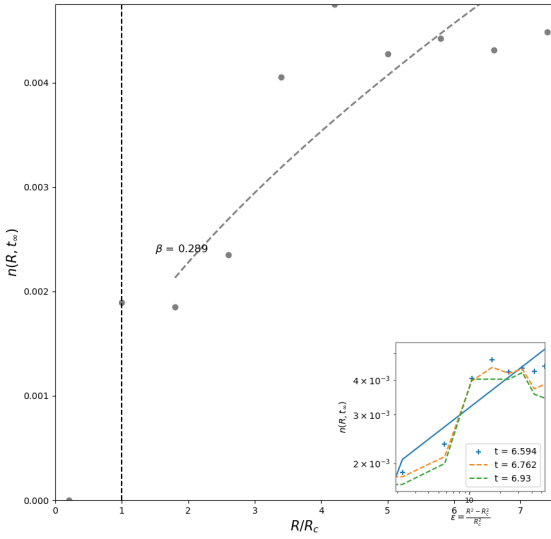
Figure A.36: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.779$. In A.36a is represented the number of people trapped in the road network at time t . In A.36b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.36c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.048$. In A.36d we see the linear increase of τ compared to Γ



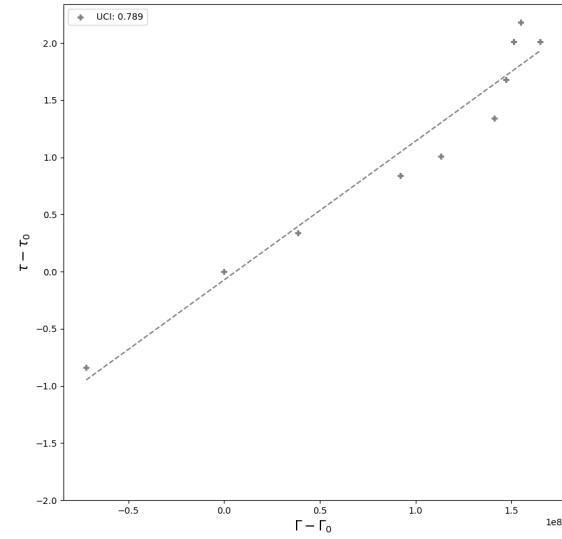
(a)



(b)

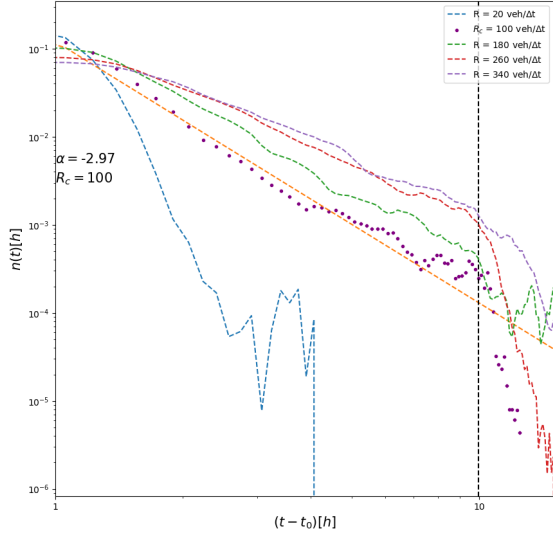


(c)

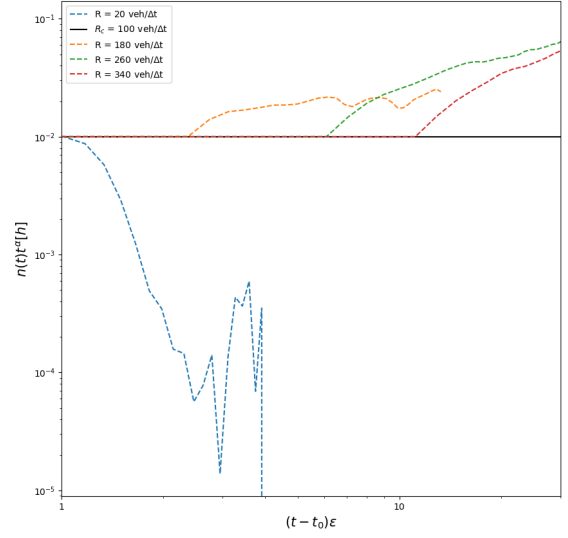


(d)

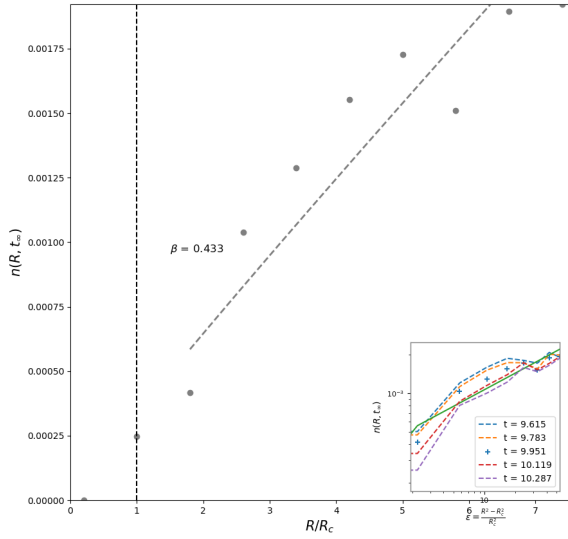
Figure A.37: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.789$. In A.37a is represented the number of people trapped in the road network at time t . In A.37b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.37c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.289$. In A.37d we see the linear increase of τ compared to Γ



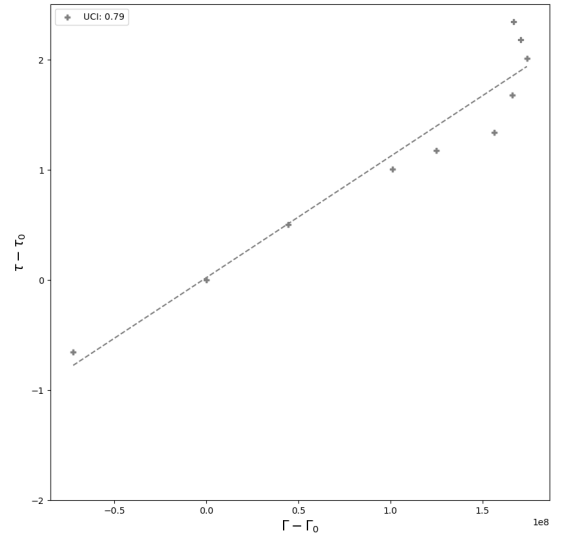
(a)



(b)

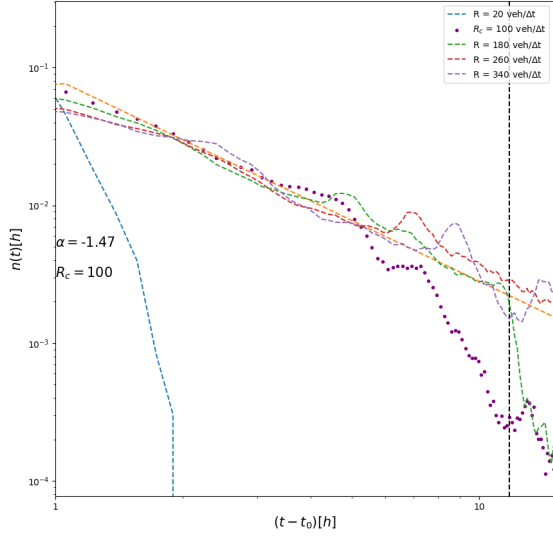


(c)

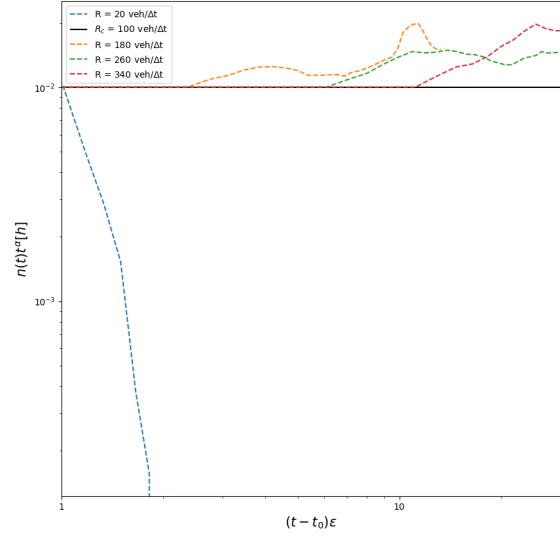


(d)

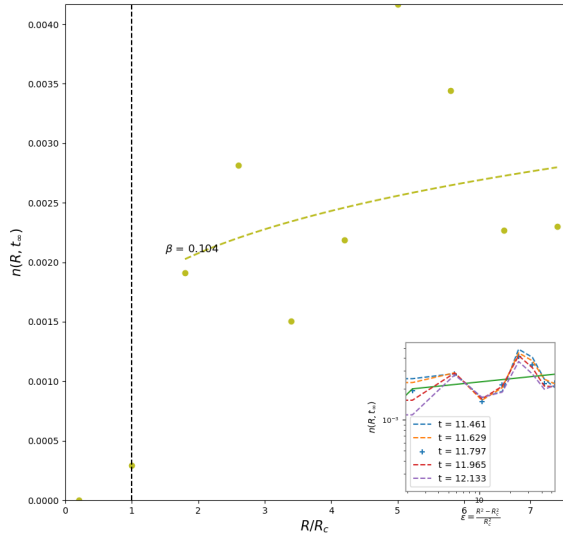
Figure A.38: Here are represented the important features about the phase transition in Rio De Janeiro for UCI = 0.79. In A.38a is represented the number of people trapped in the road network at time t . In A.38b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.38c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.433$. In A.38d we see the linear increase of τ compared to Γ



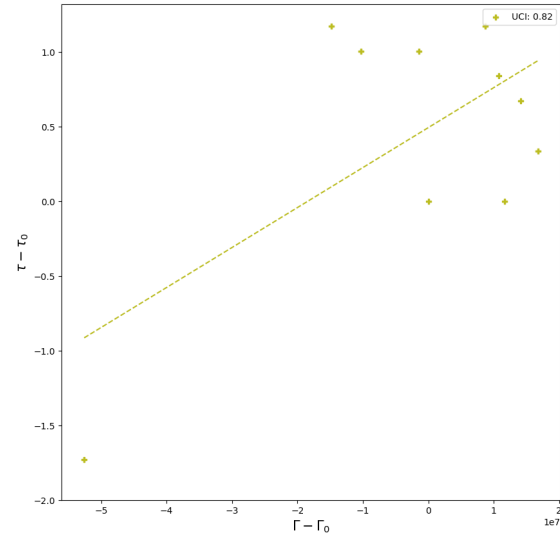
(a)



(b)

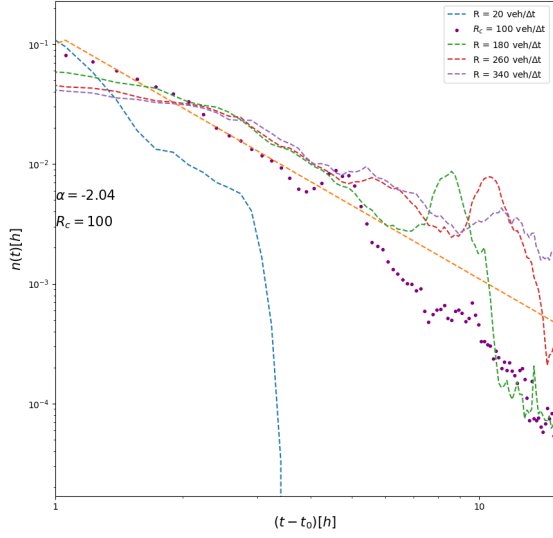


(c)

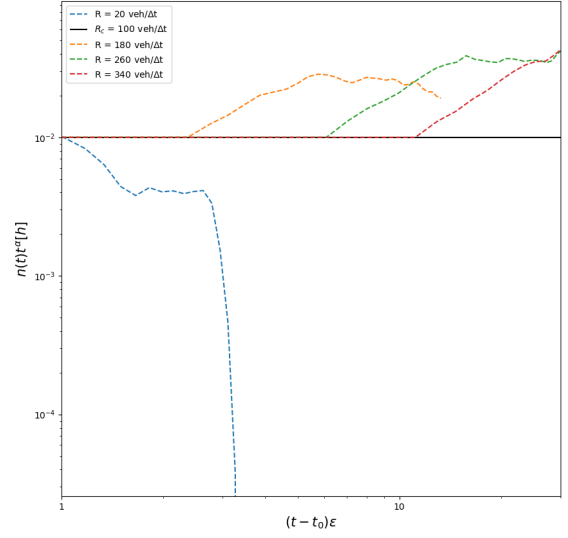


(d)

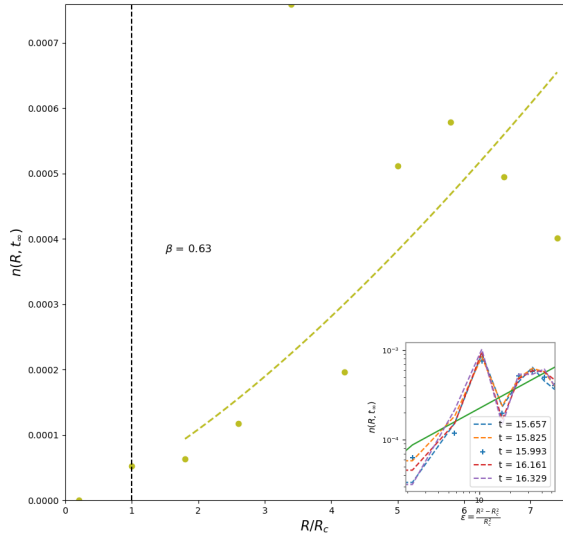
Figure A.39: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.82$. In A.39a is represented the number of people trapped in the road network at time t . In A.39b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.39c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.104$. In A.39d we see the linear increase of τ compared to Γ



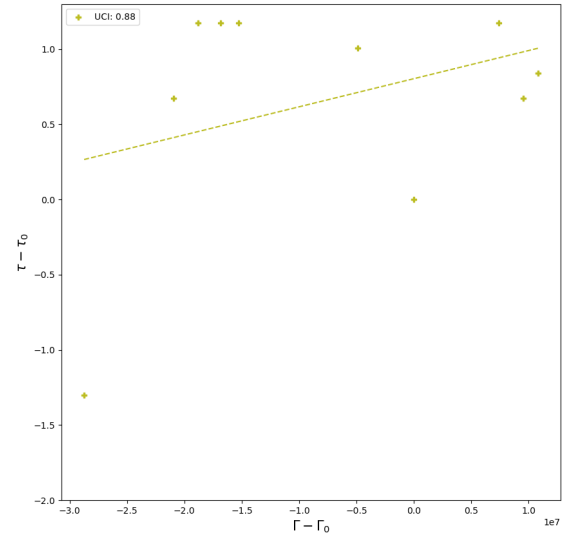
(a)



(b)

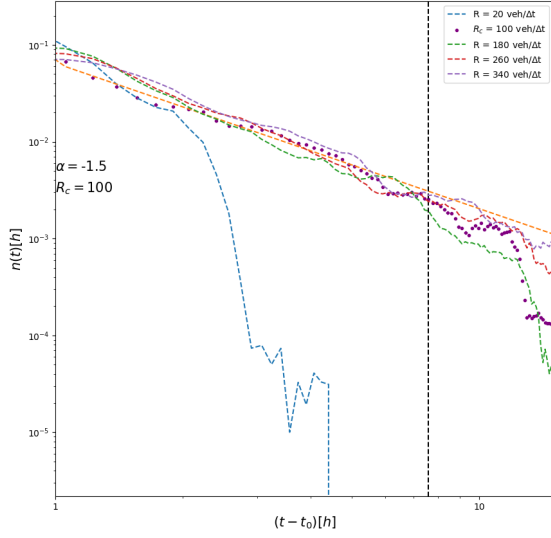


(c)

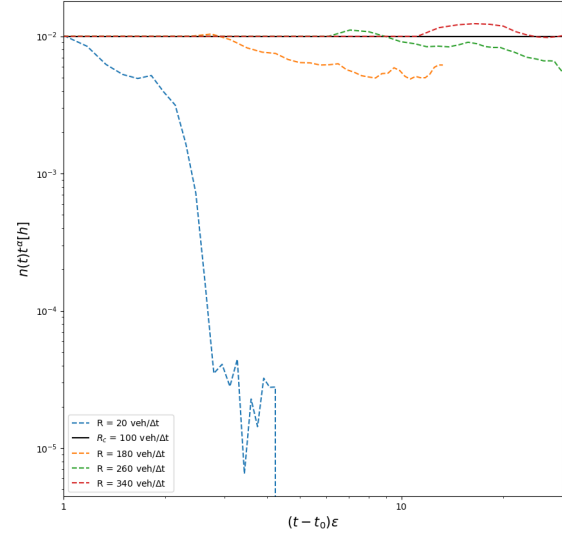


(d)

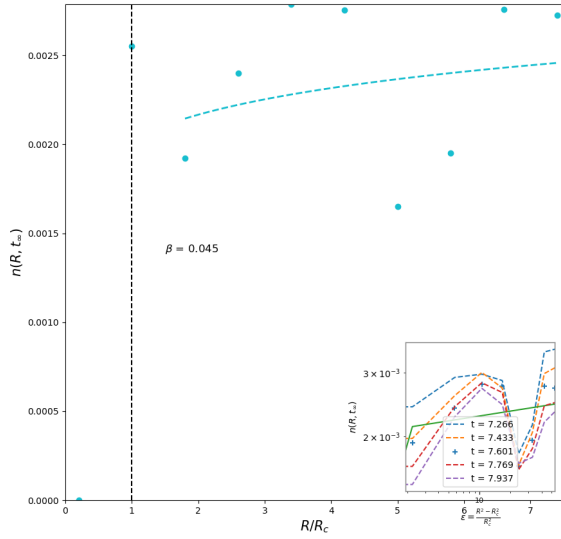
Figure A.40: Here are represented the important features about the phase transition in Rio De Janeiro for UCI = 0.88. In A.40a is represented the number of people trapped in the road network at time t . In A.40b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.40c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.630$. In A.40d we see the linear increase of τ compared to Γ



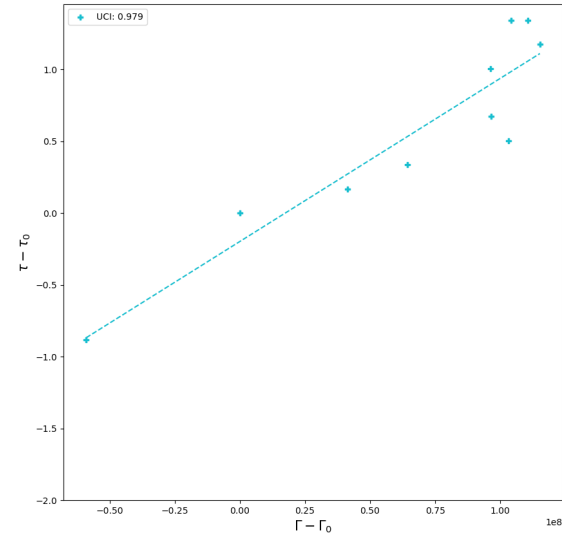
(a)



(b)

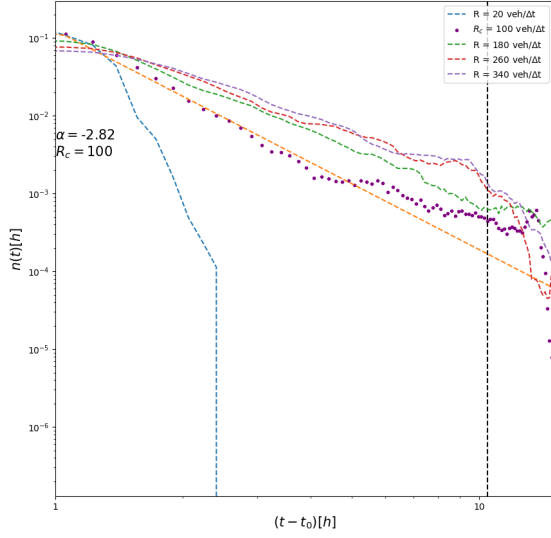


(c)

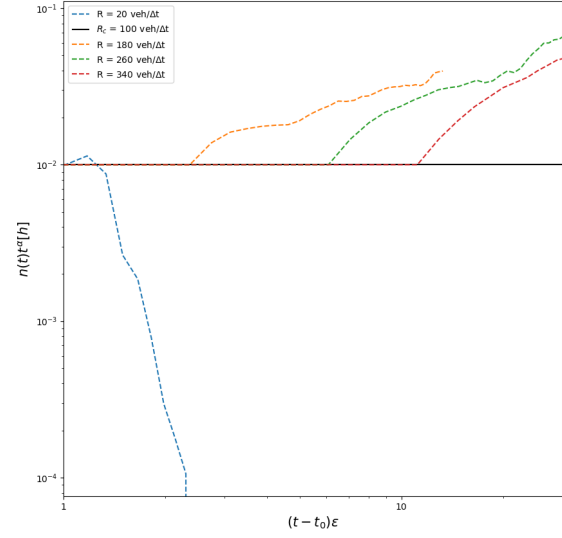


(d)

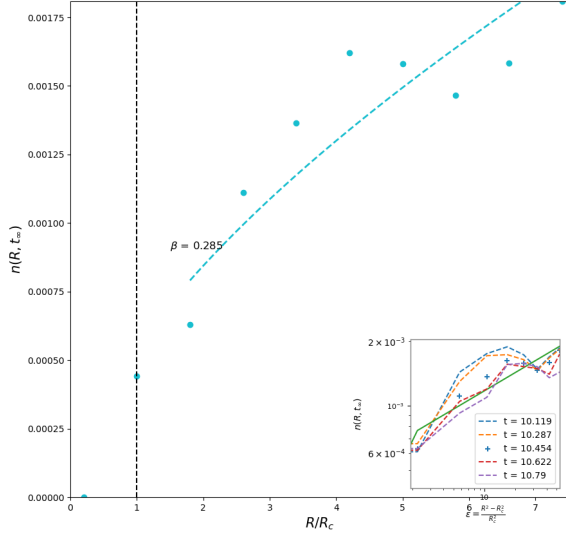
Figure A.41: Here are represented the important features about the phase transition in Rio De Janeiro for $UCI = 0.979$. In A.41a is represented the number of people trapped in the road network at time t . In A.41b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.41c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.045$. In A.41d we see the linear increase of τ compared to Γ



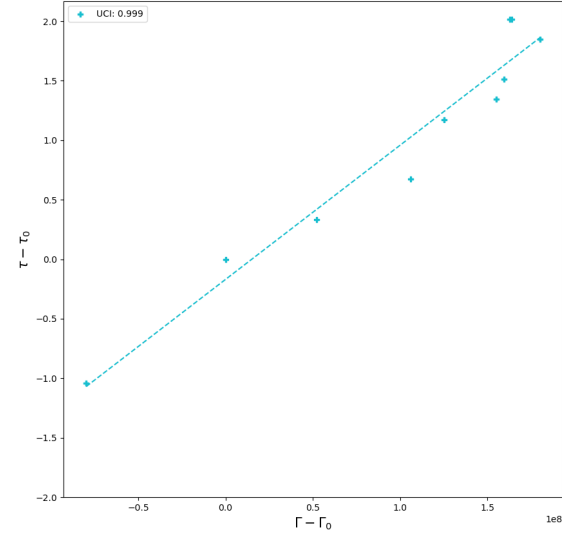
(a)



(b)

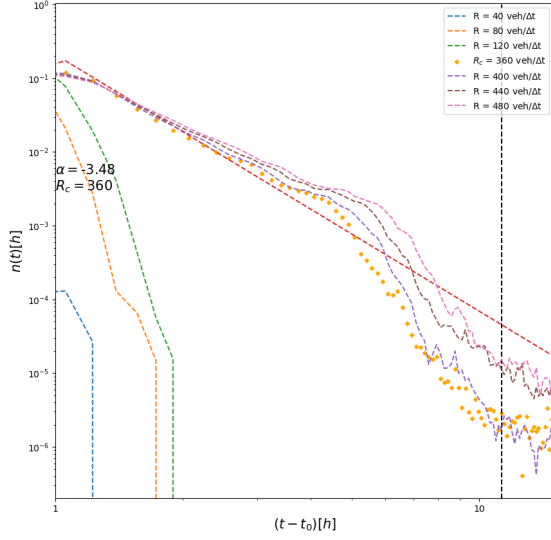


(c)

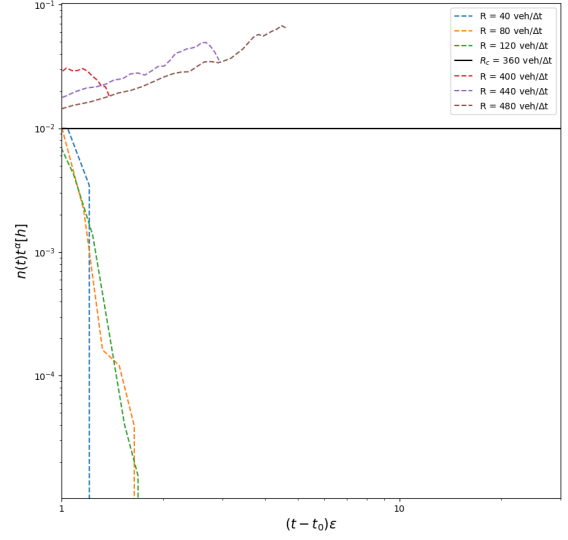


(d)

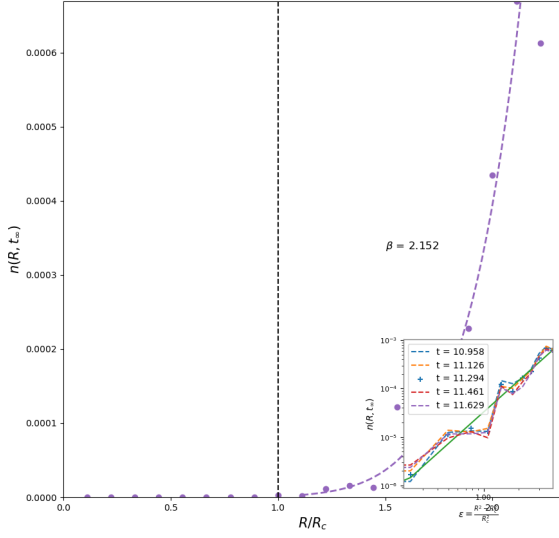
Figure A.42: Here are represented the important features about the phase transition in Rio De Janeiro for UCI = 0.999. In A.42a is represented the number of people trapped in the road network at time t . In A.42b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.42c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.285$. In A.42d we see the linear increase of τ compared to Γ



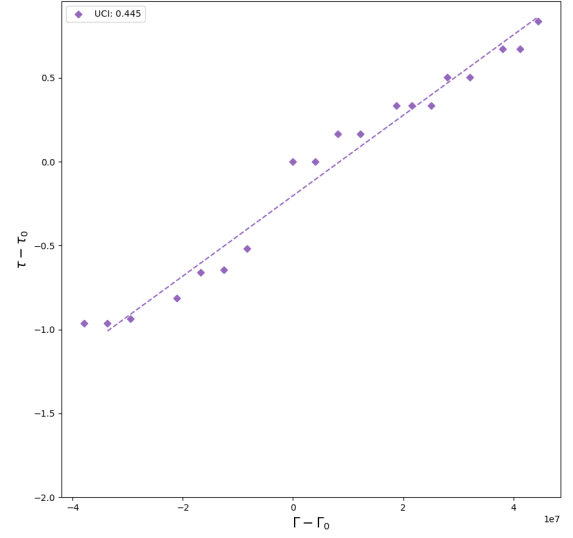
(a)



(b)

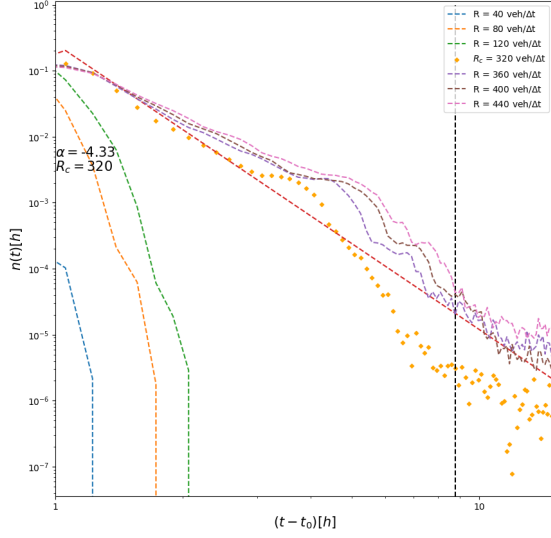


(c)

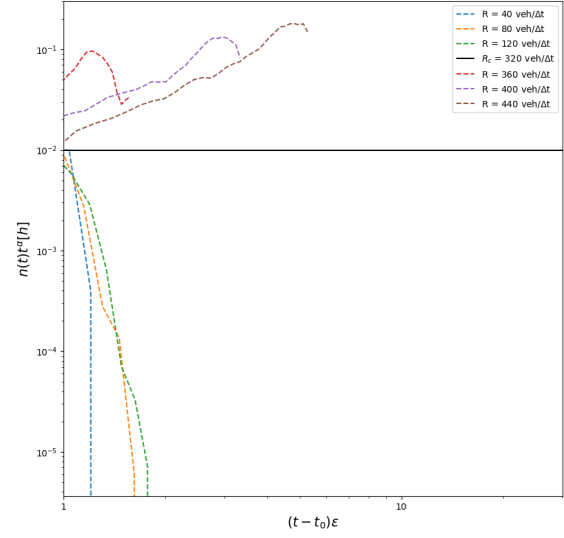


(d)

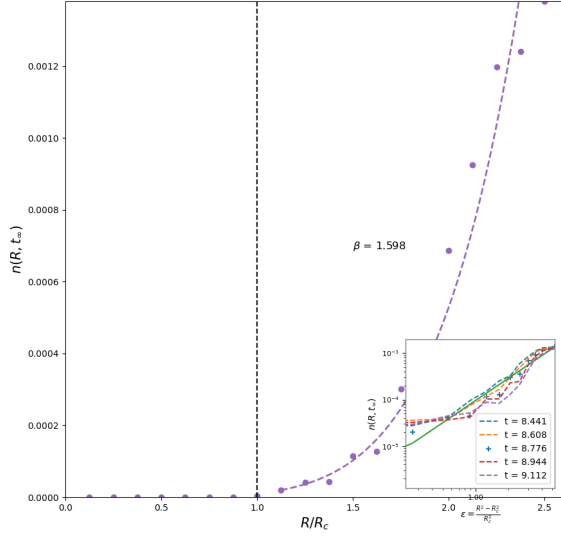
Figure A.43: Here are represented the important features about the phase transition in San Francisco for UCI = 0.445. In A.43a is represented the number of people trapped in the road network at time t . In A.43b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.43c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 2.152$. In A.43d we see the linear increase of τ compared to Γ



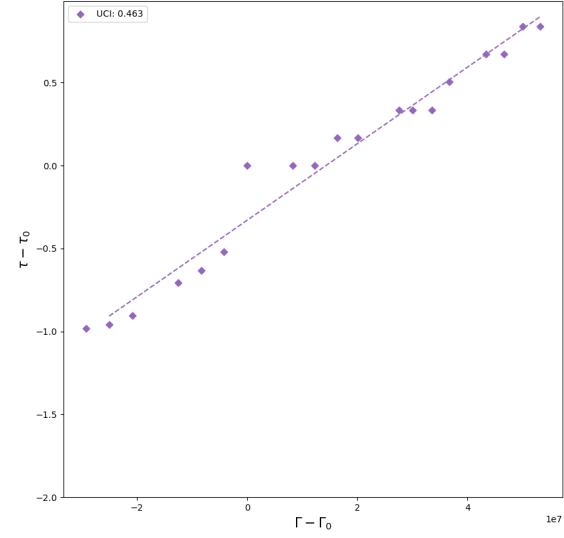
(a)



(b)

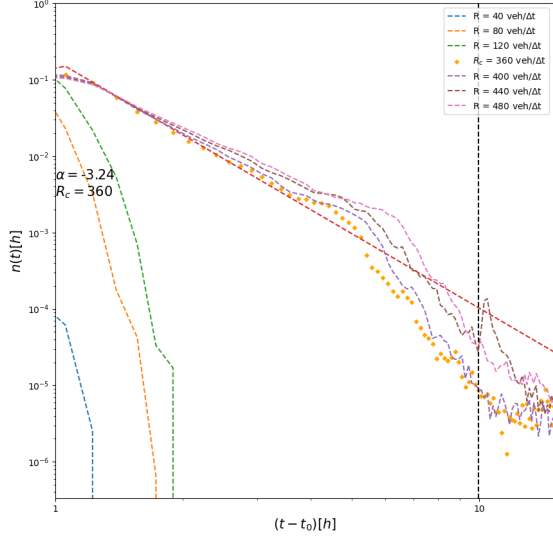


(c)

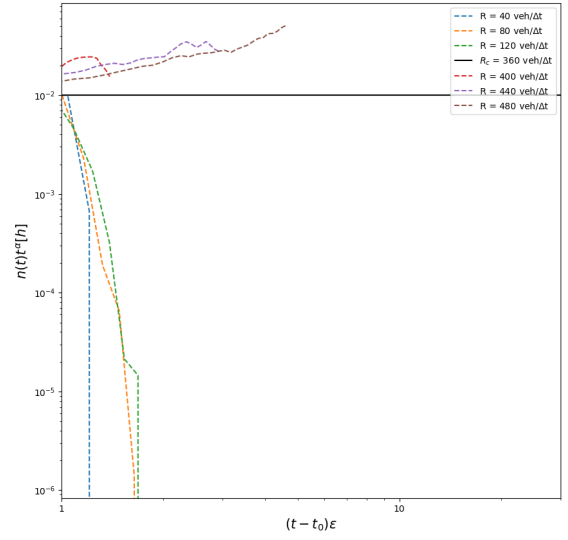


(d)

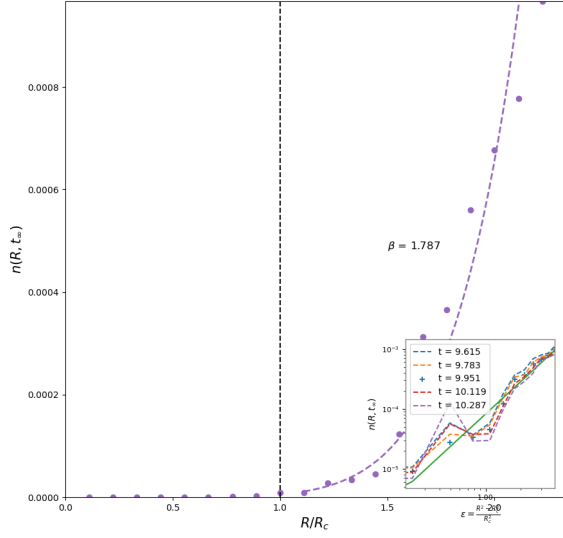
Figure A.44: Here are represented the important features about the phase transition in San Francisco for UCI = 0.463. In A.44a is represented the number of people trapped in the road network at time t . In A.44b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.44c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.598$. In A.44d we see the linear increase of τ compared to Γ



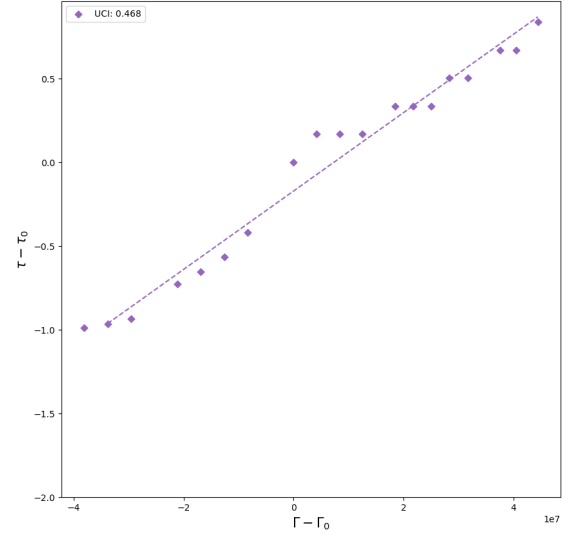
(a)



(b)

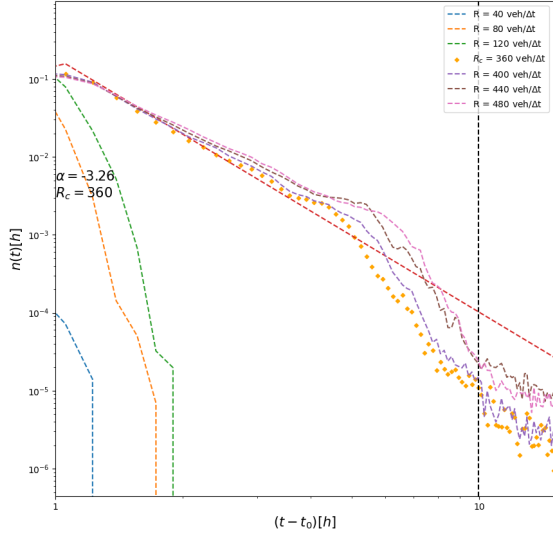


(c)

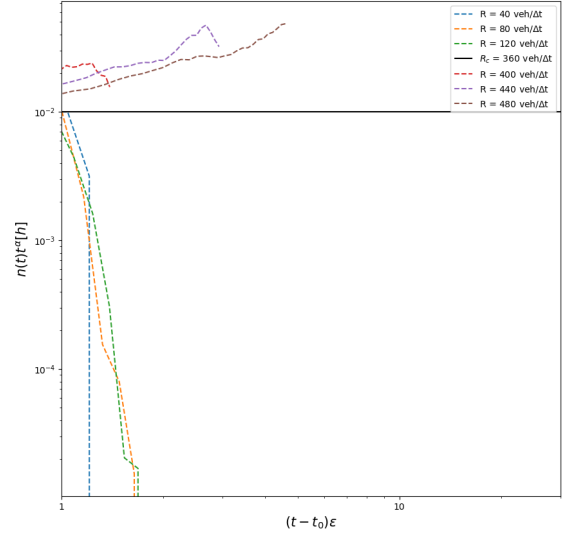


(d)

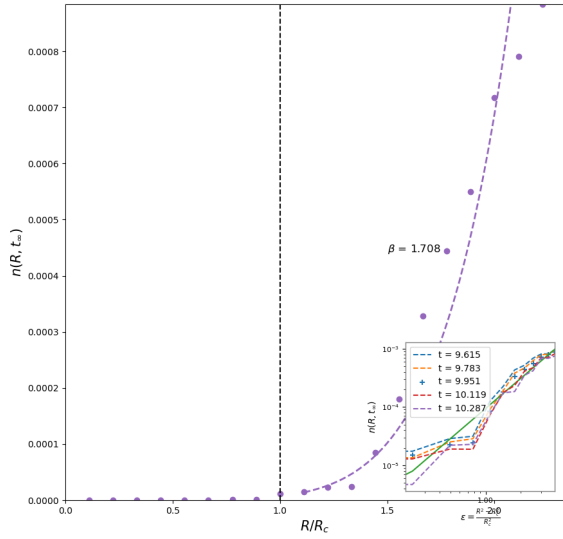
Figure A.45: Here are represented the important features about the phase transition in San Francisco for UCI = 0.468. In A.45a is represented the number of people trapped in the road network at time t . In A.45b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.45c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.787$. In A.45d we see the linear increase of τ compared to Γ



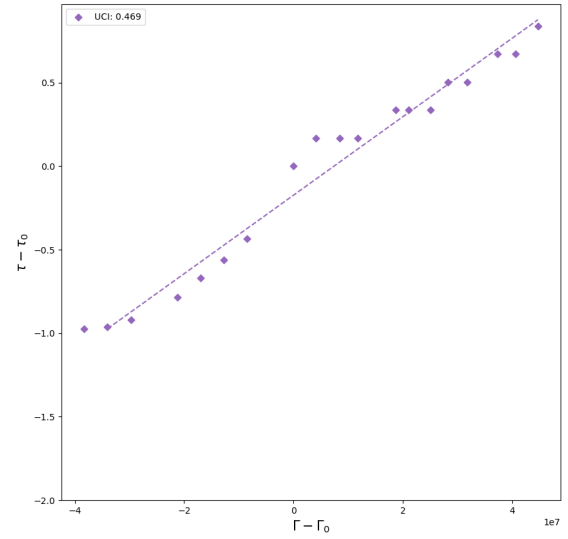
(a)



(b)



(c)



(d)

Figure A.46: Here are represented the important features about the phase transition in San Francisco for UCI = 0.469. In A.46a is represented the number of people trapped in the road network at time t . In A.46b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.46c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.708$. In A.46d we see the linear increase of τ compared to Γ

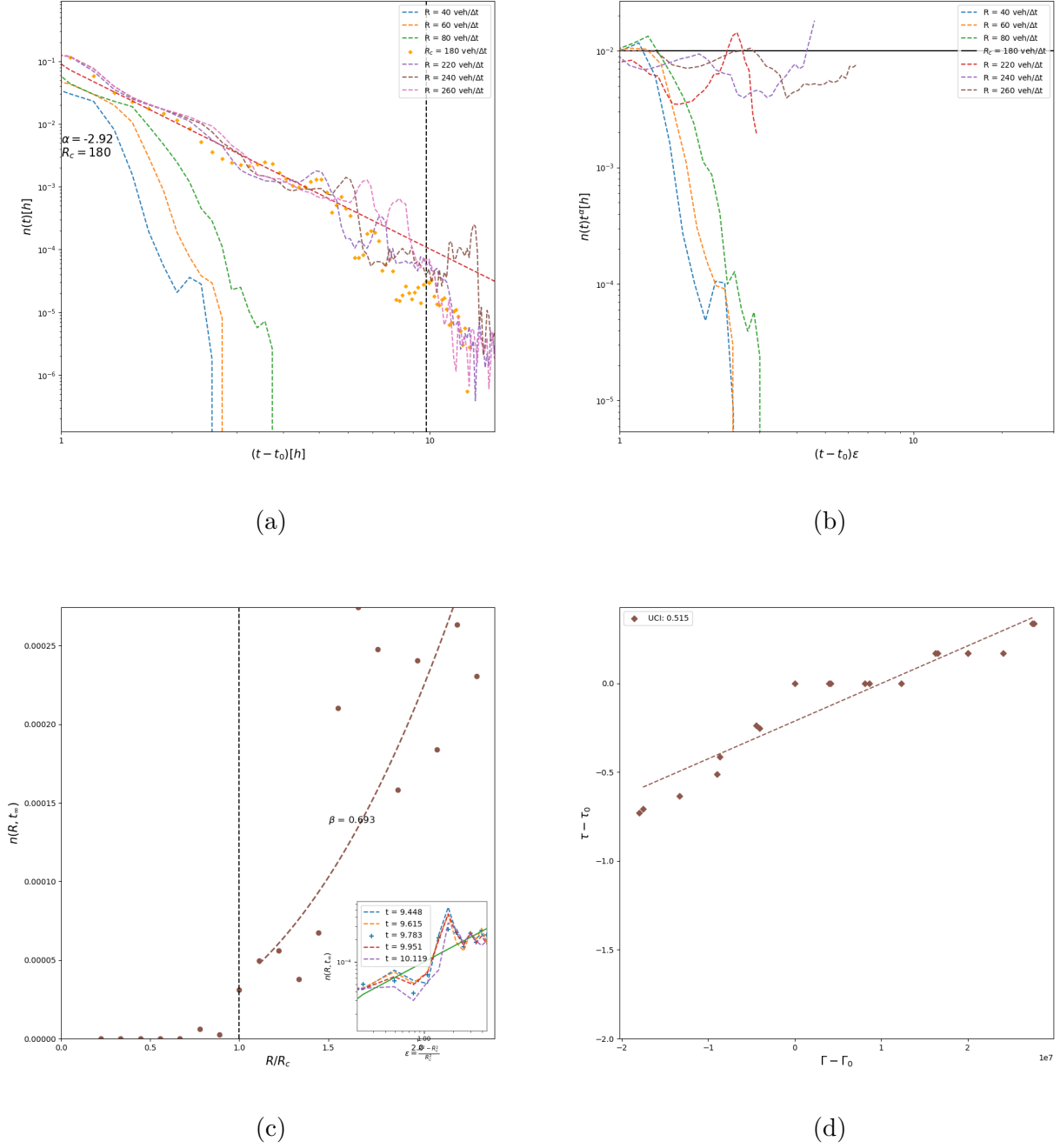
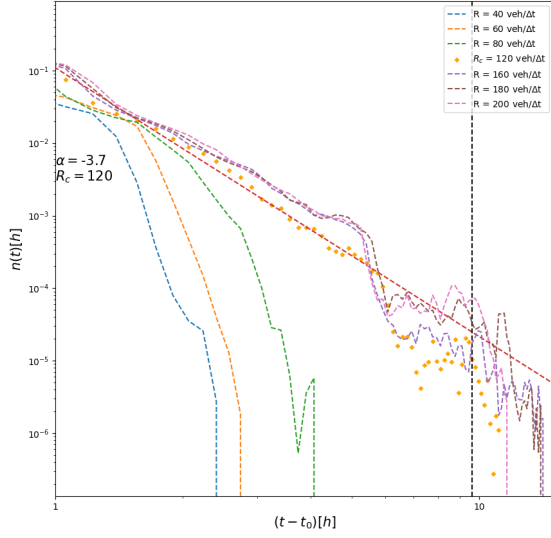
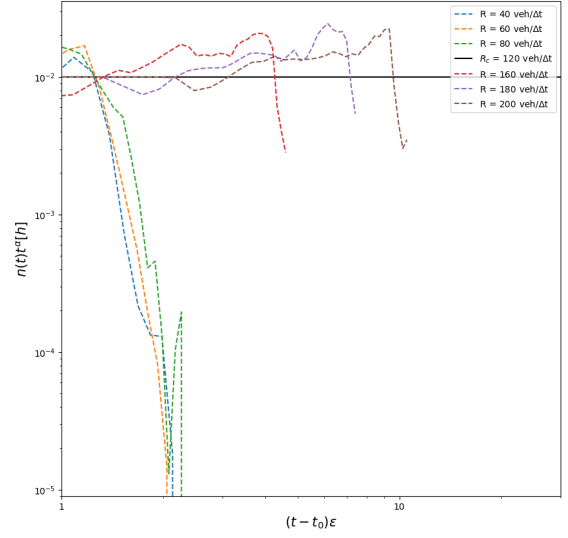


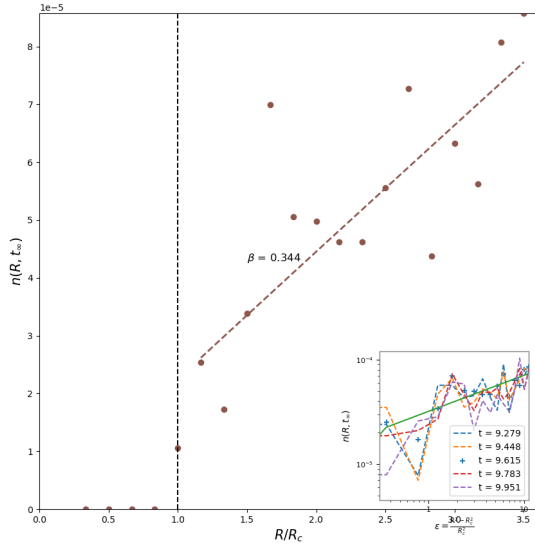
Figure A.47: Here are represented the important features about the phase transition in San Francisco for UCI = 0.515. In A.47a is represented the number of people trapped in the road network at time t . In A.47b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.47c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.693$. In A.47d we see the linear increase of τ compared to Γ



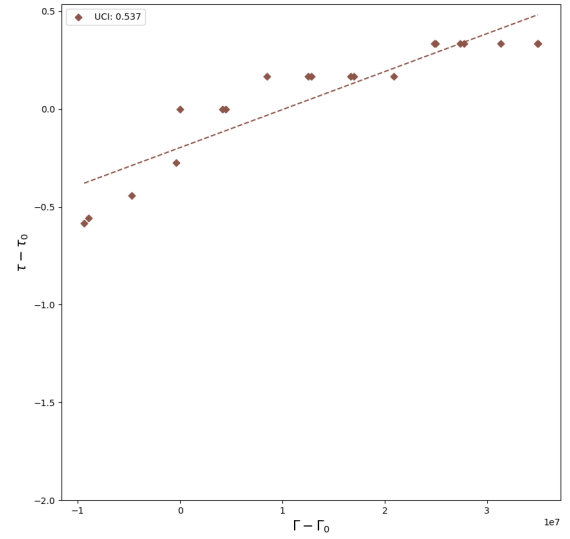
(a)



(b)

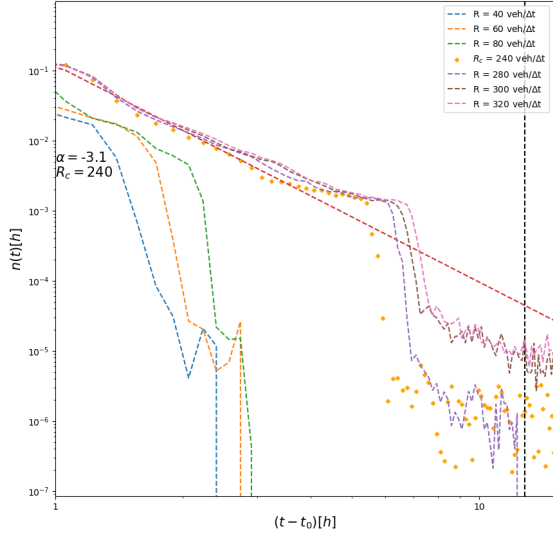


(c)

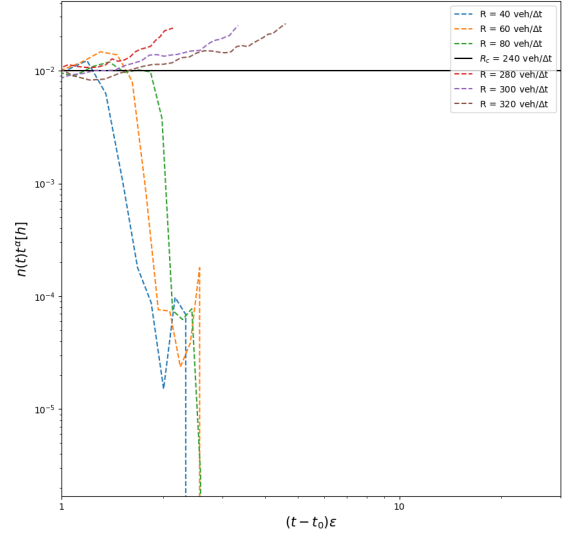


(d)

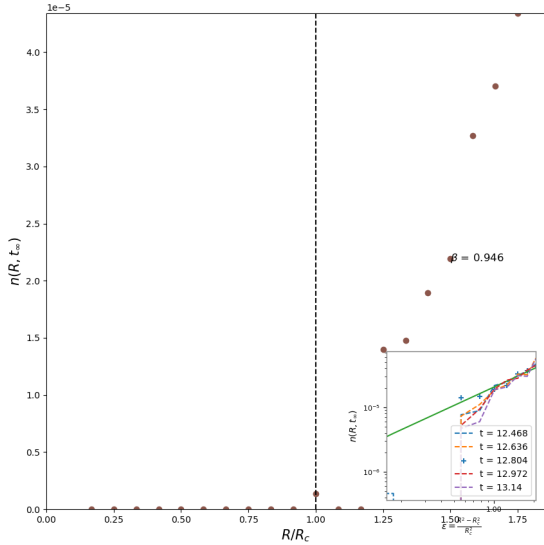
Figure A.48: Here are represented the important features about the phase transition in San Francisco for UCI = 0.537. In A.48a is represented the number of people trapped in the road network at time t . In A.48b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.48c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.344$. In A.48d we see the linear increase of τ compared to Γ



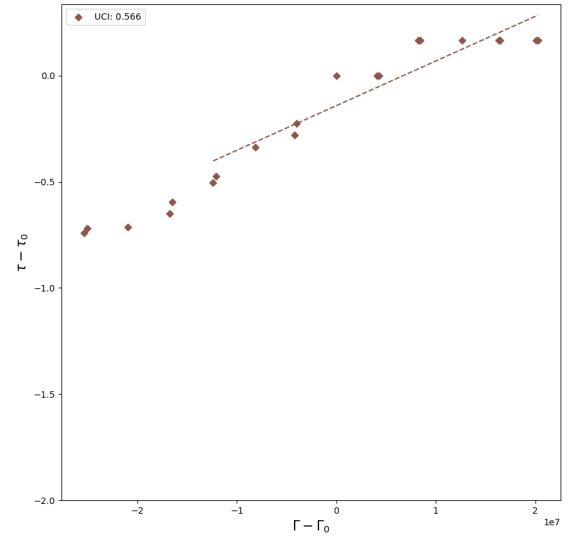
(a)



(b)



(c)



(d)

Figure A.49: Here are represented the important features about the phase transition in San Francisco for UCI = 0.566. In A.49a is represented the number of people trapped in the road network at time t . In A.49b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.49c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.946$. In A.49d we see the linear increase of τ compared to Γ

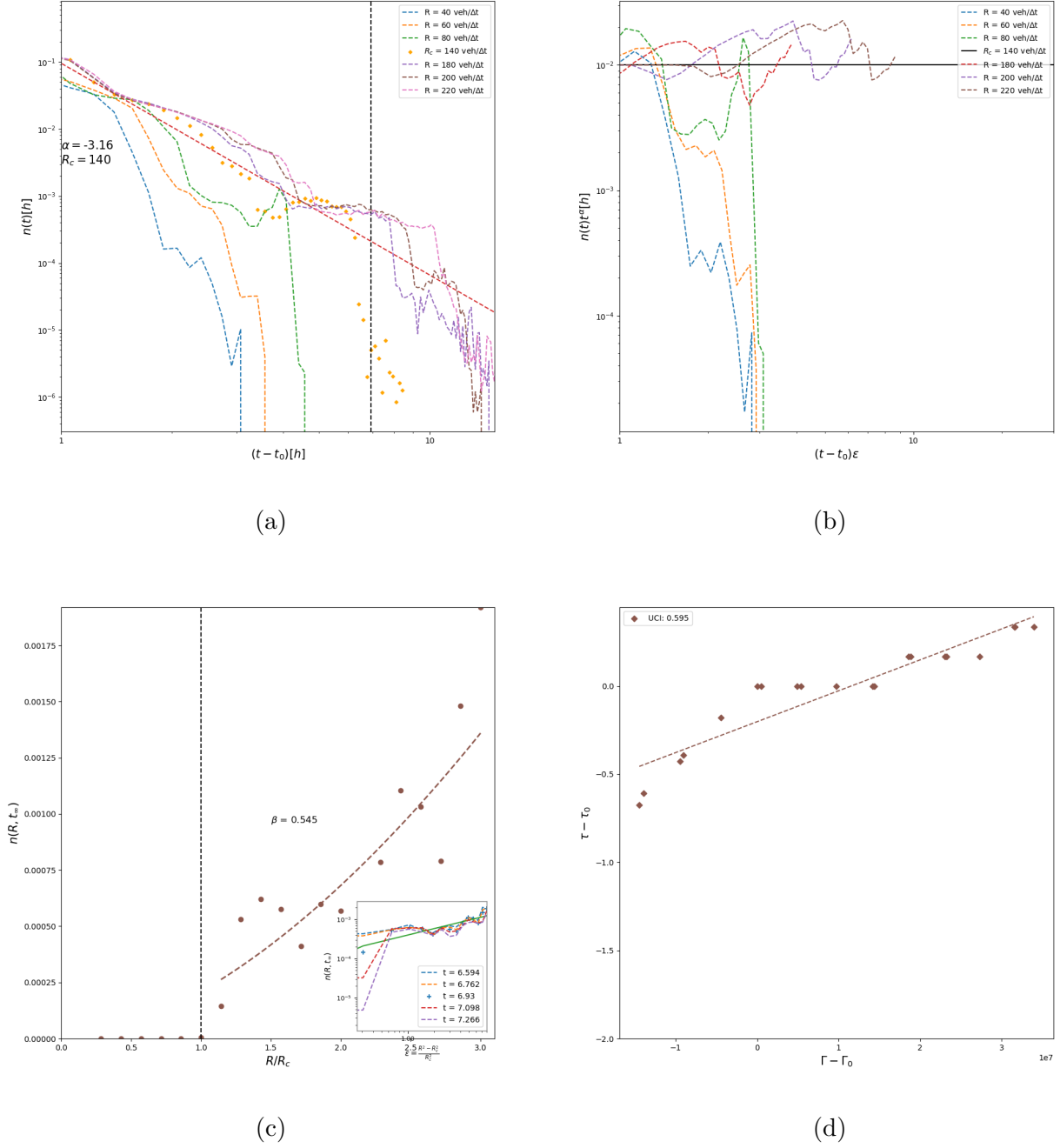
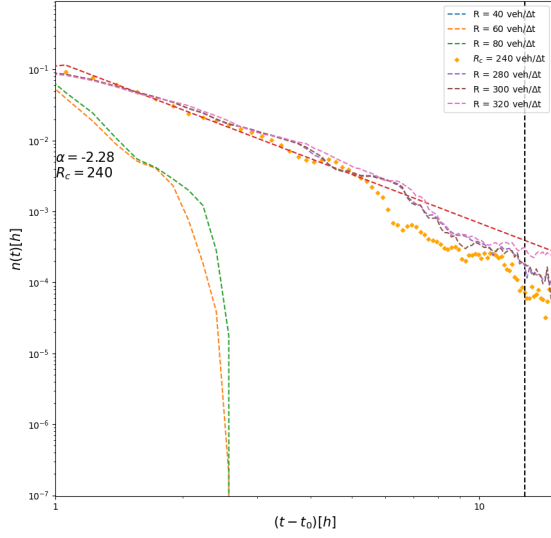
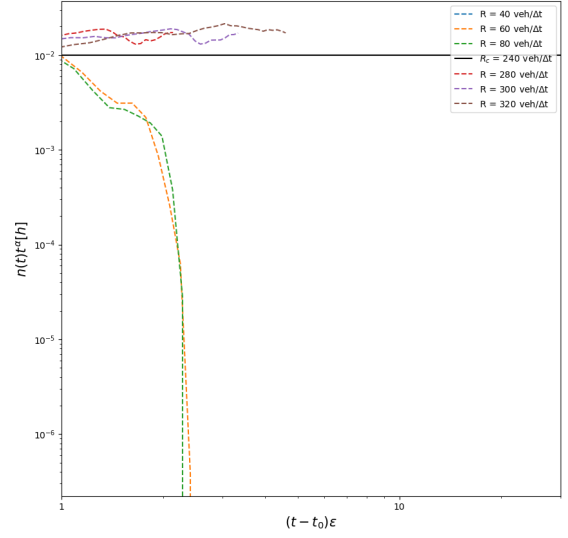


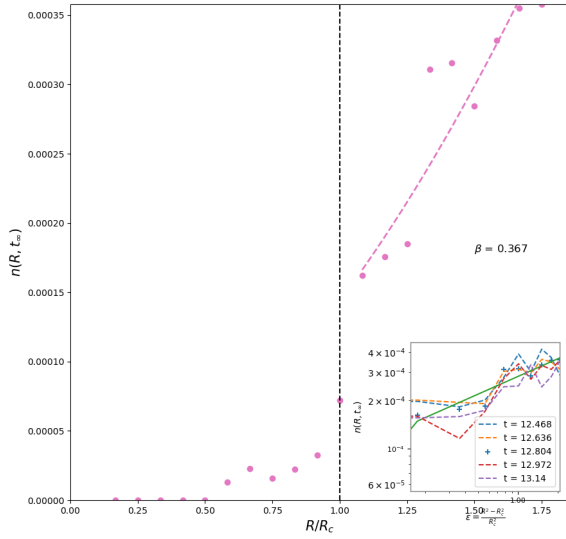
Figure A.50: Here are represented the important features about the phase transition in San Francisco for UCI = 0.595. In A.50a is represented the number of people trapped in the road network at time t . In A.50b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.50c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.545$. In A.50d we see the linear increase of τ compared to Γ



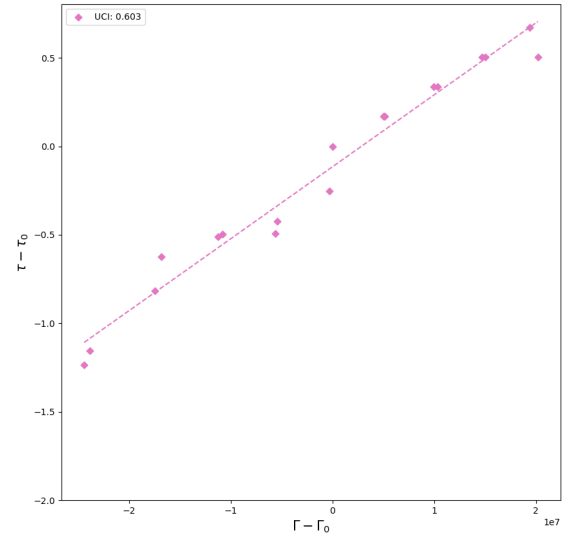
(a)



(b)

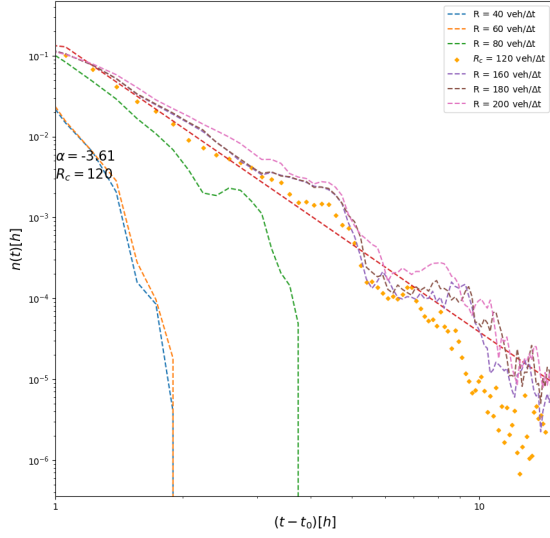


(c)

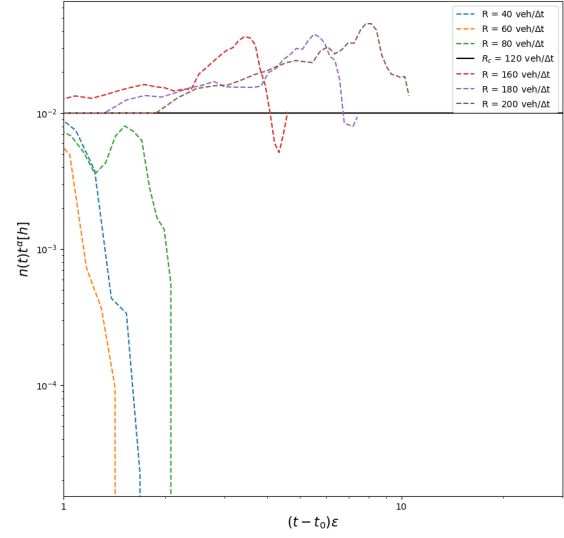


(d)

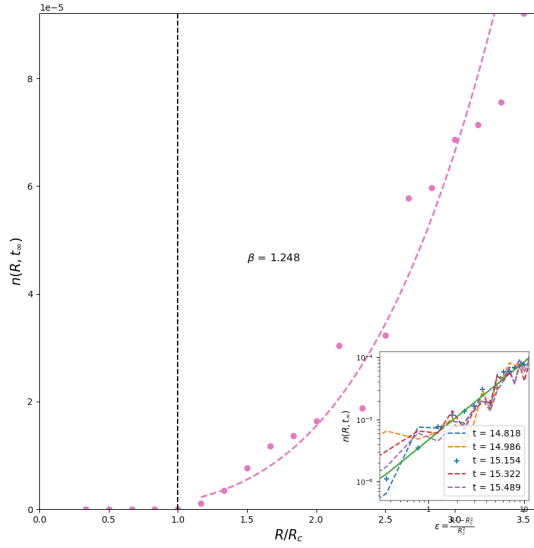
Figure A.51: Here are represented the important features about the phase transition in San Francisco for UCI = 0.603. In A.51a is represented the number of people trapped in the road network at time t . In A.51b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.51c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.367$. In A.51d we see the linear increase of τ compared to Γ



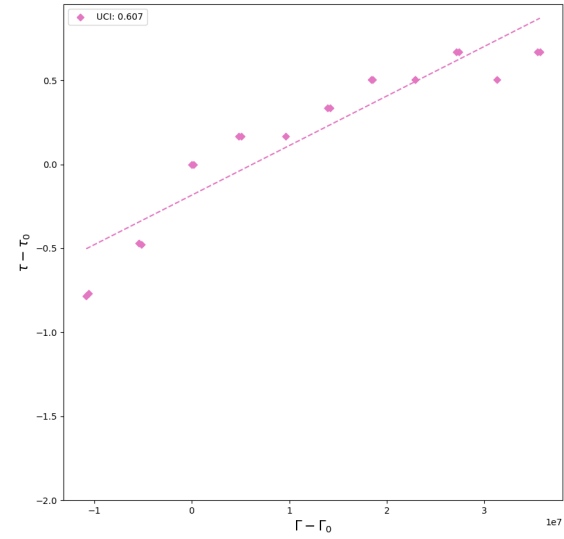
(a)



(b)

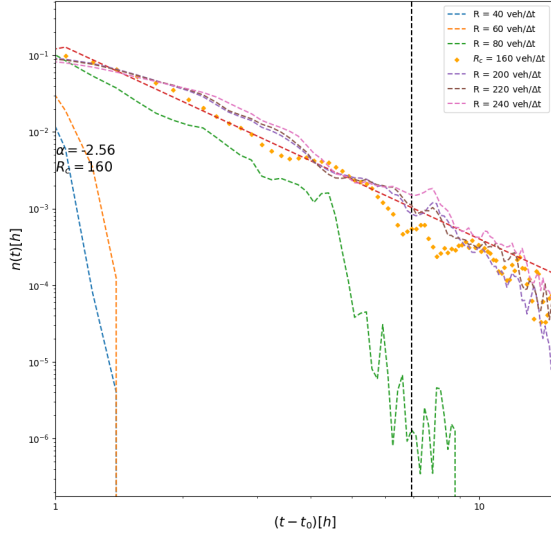


(c)

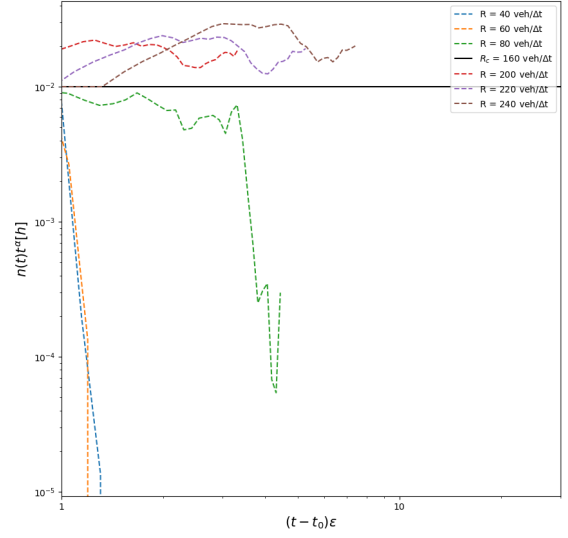


(d)

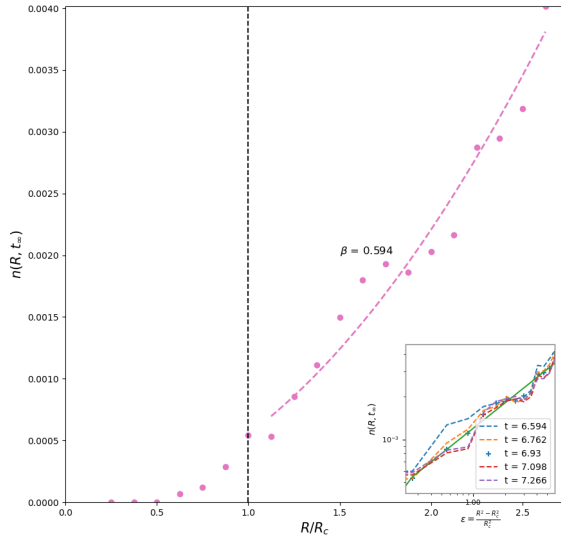
Figure A.52: Here are represented the important features about the phase transition in San Francisco for UCI = 0.607. In A.52a is represented the number of people trapped in the road network at time t . In A.52b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.52c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.248$. In A.52d we see the linear increase of τ compared to Γ



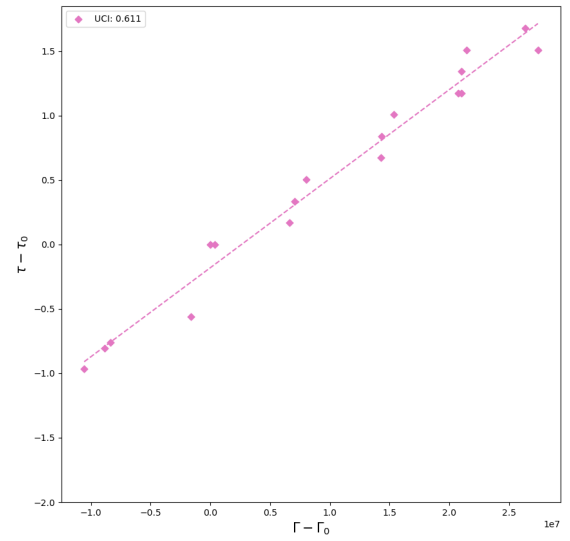
(a)



(b)

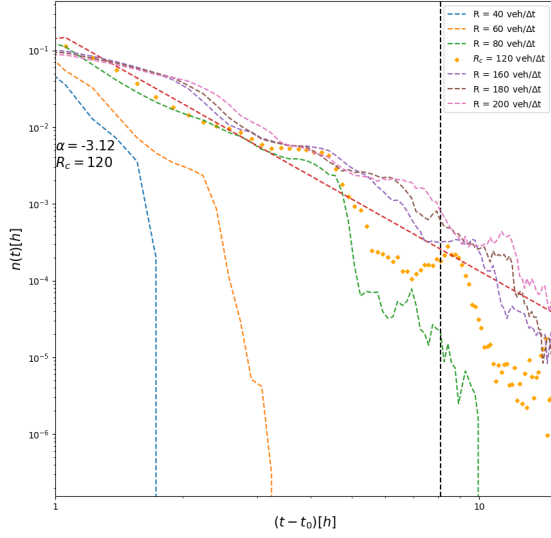


(c)

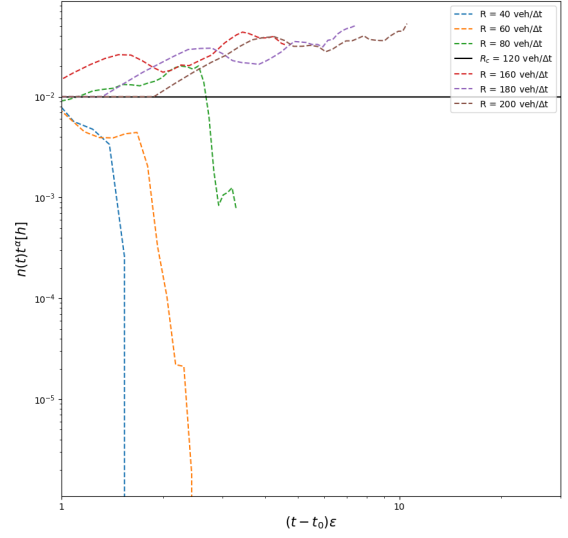


(d)

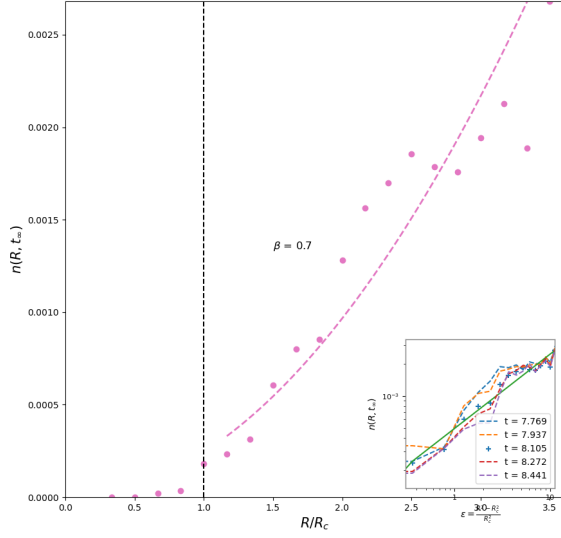
Figure A.53: Here are represented the important features about the phase transition in San Francisco for UCI = 0.611. In A.53a is represented the number of people trapped in the road network at time t . In A.53b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.53c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.594$. In A.53d we see the linear increase of τ compared to Γ



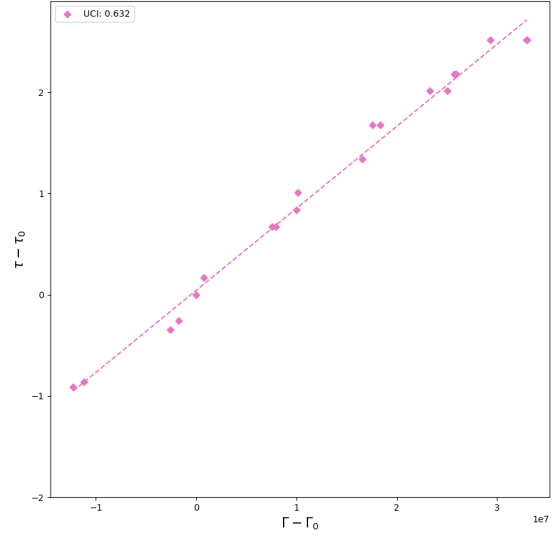
(a)



(b)

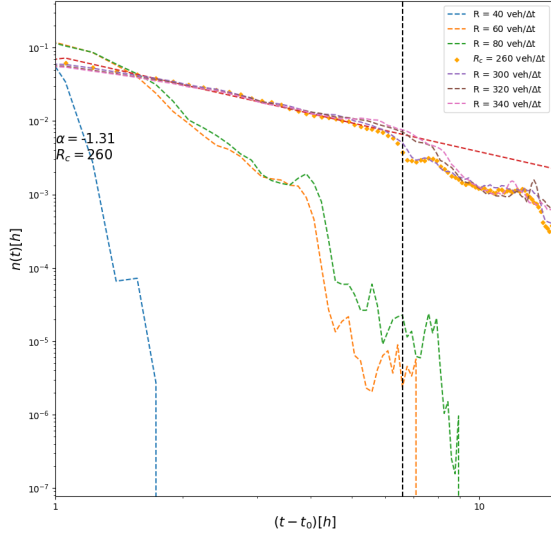


(c)

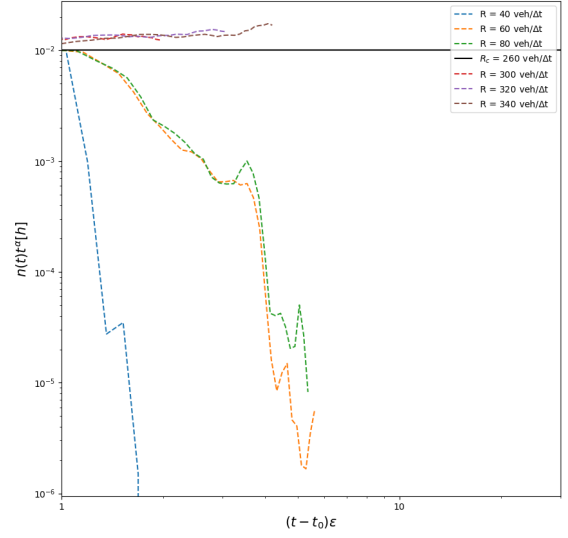


(d)

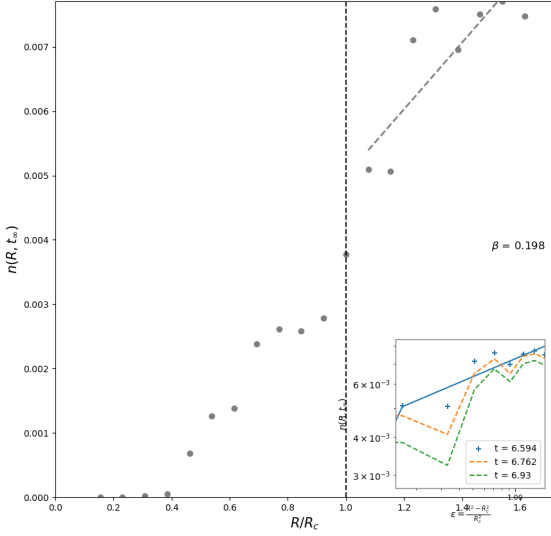
Figure A.54: Here are represented the important features about the phase transition in San Francisco for UCI = 0.632. In A.54a is represented the number of people trapped in the road network at time t . In A.54b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.54c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.700$. In A.54d we see the linear increase of τ compared to Γ



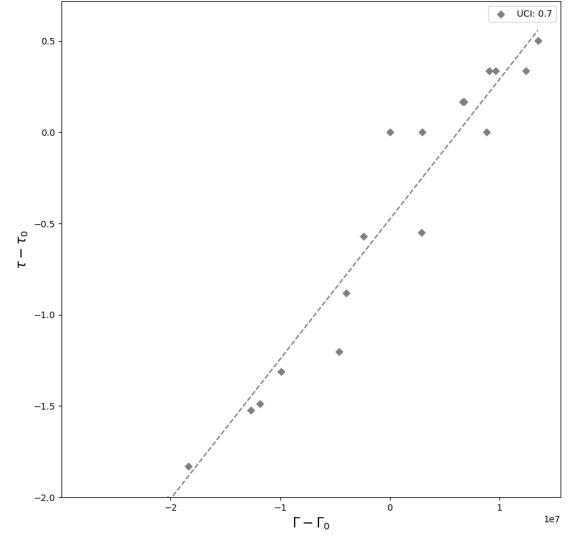
(a)



(b)

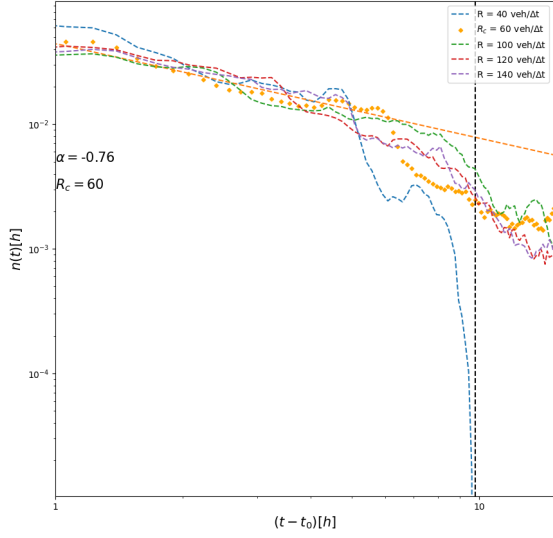


(c)

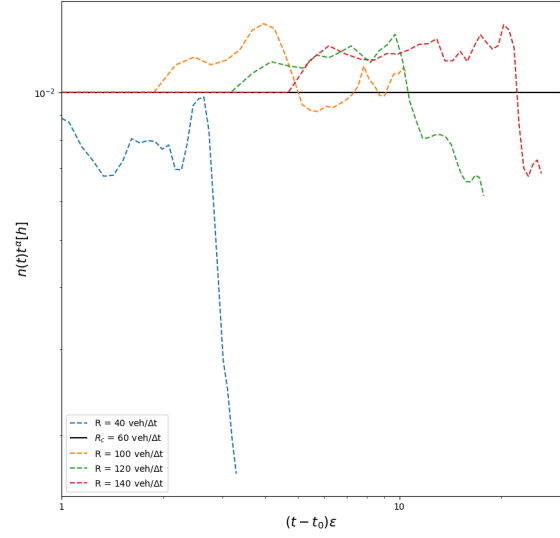


(d)

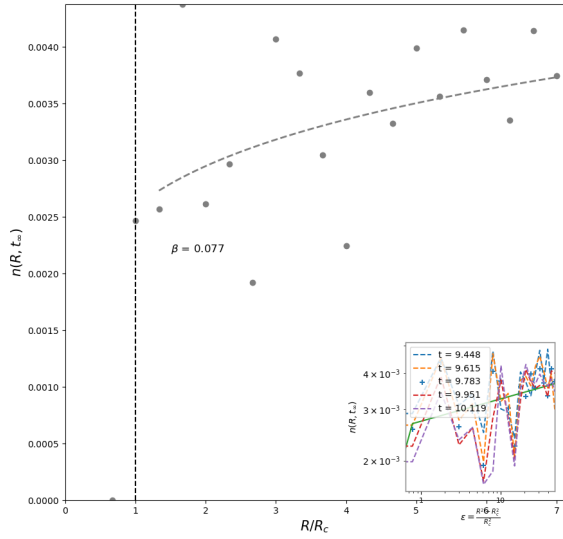
Figure A.55: Here are represented the important features about the phase transition in San Francisco for $UCI = 0.7$. In A.55a is represented the number of people trapped in the road network at time t . In A.55b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.55c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.198$. In A.55d we see the linear increase of τ compared to Γ



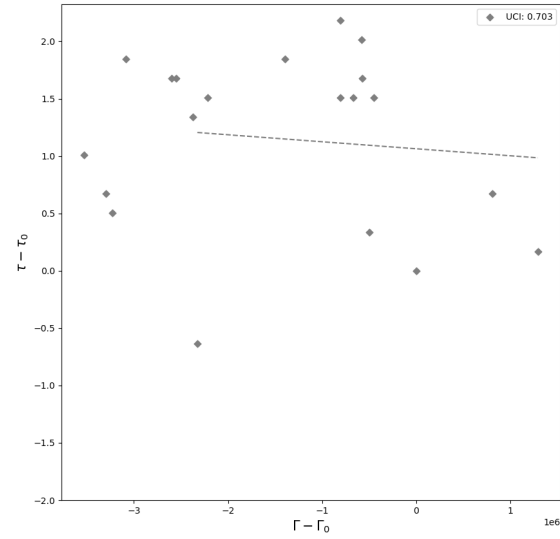
(a)



(b)

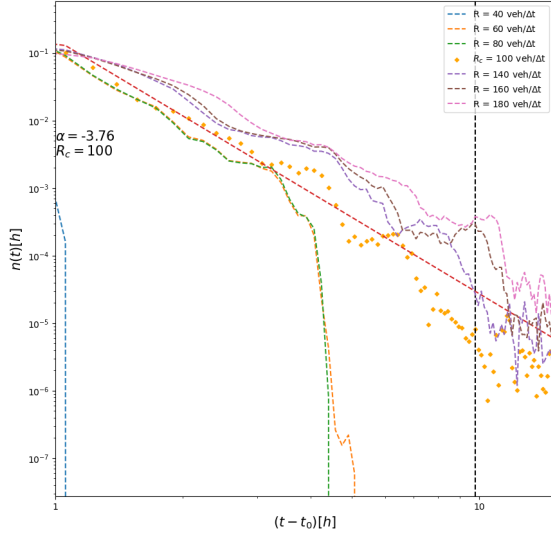


(c)

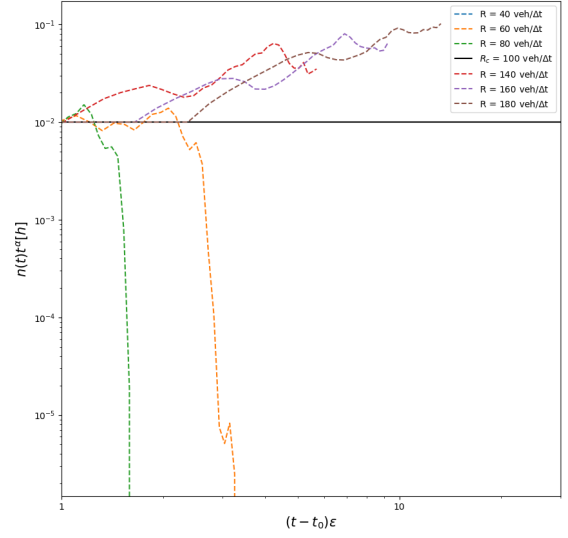


(d)

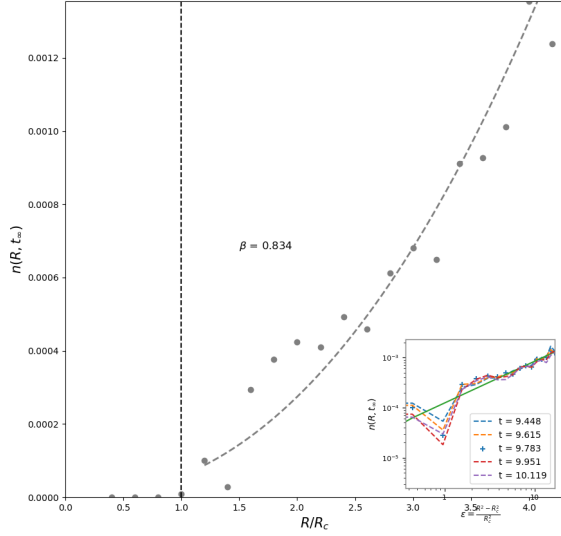
Figure A.56: Here are represented the important features about the phase transition in San Francisco for UCI = 0.703. In A.56a is represented the number of people trapped in the road network at time t . In A.56b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.56c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.077$. In A.56d we see the linear increase of τ compared to Γ



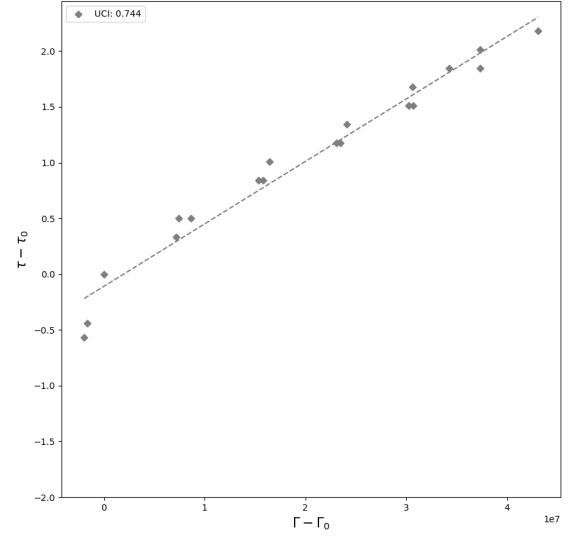
(a)



(b)

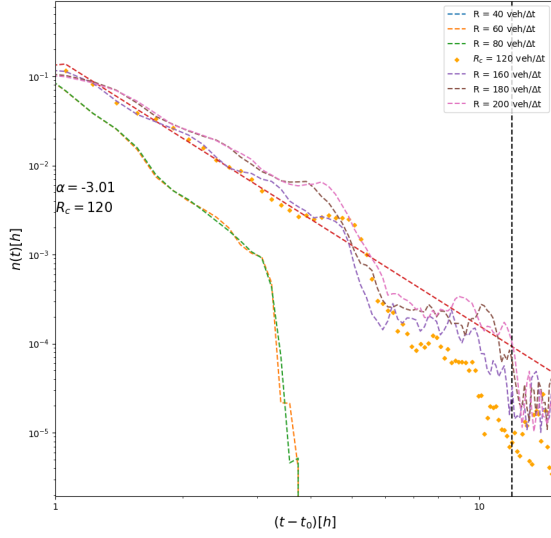


(c)

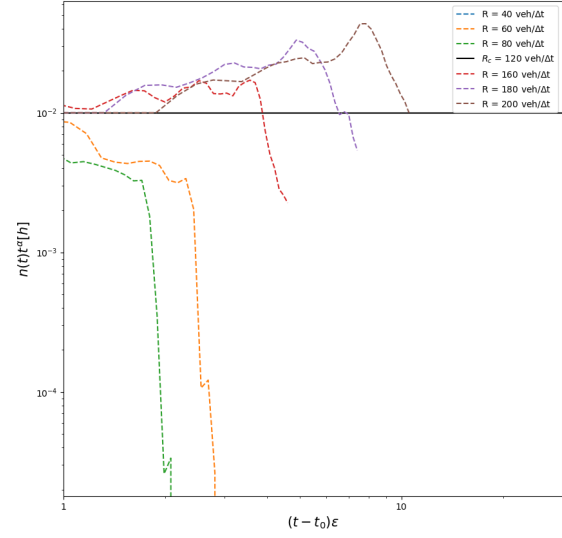


(d)

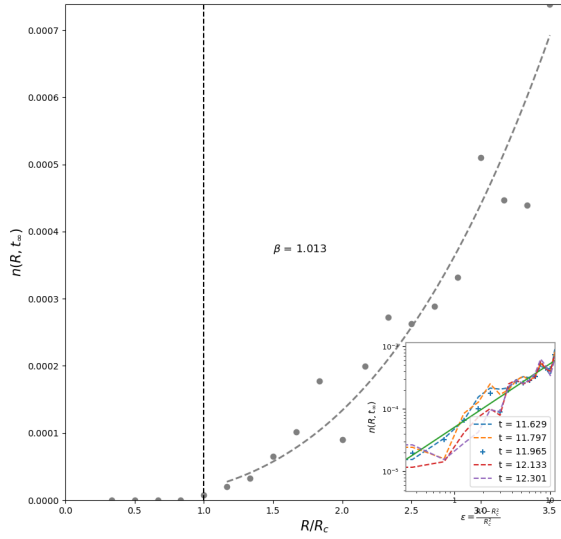
Figure A.57: Here are represented the important features about the phase transition in San Francisco for UCI = 0.744. In A.57a is represented the number of people trapped in the road network at time t . In A.57b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.57c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.834$. In A.57d we see the linear increase of τ compared to Γ



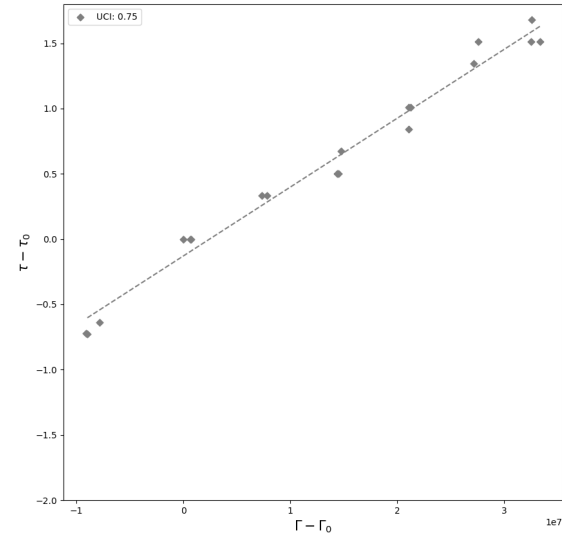
(a)



(b)

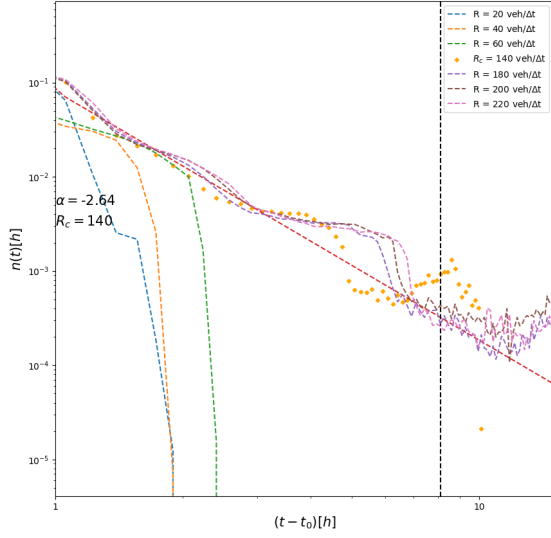


(c)

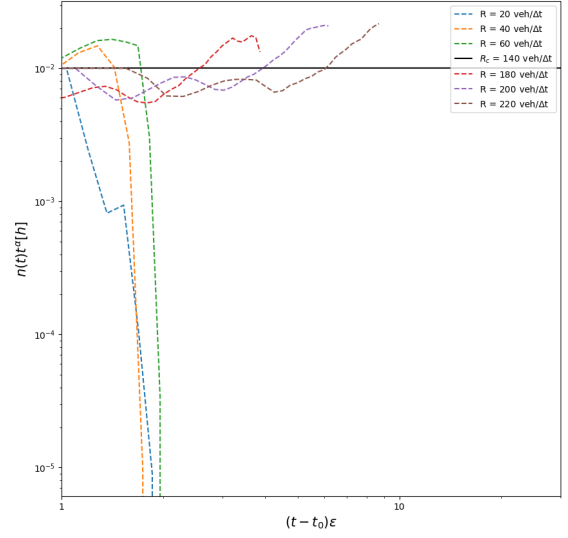


(d)

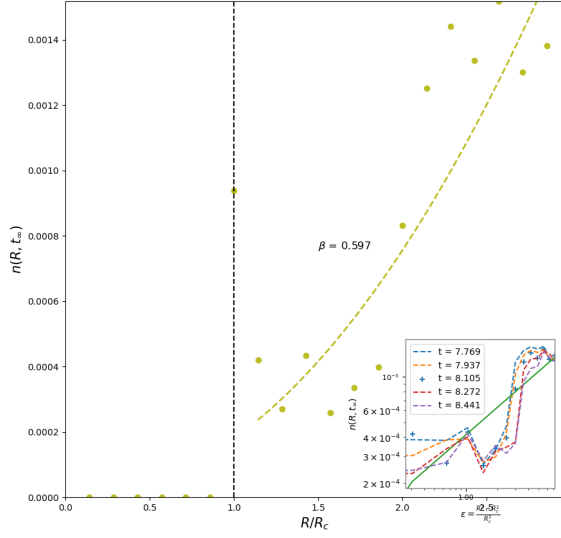
Figure A.58: Here are represented the important features about the phase transition in San Francisco for $UCI = 0.75$. In A.58a is represented the number of people trapped in the road network at time t . In A.58b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.58c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 1.013$. In A.58d we see the linear increase of τ compared to Γ



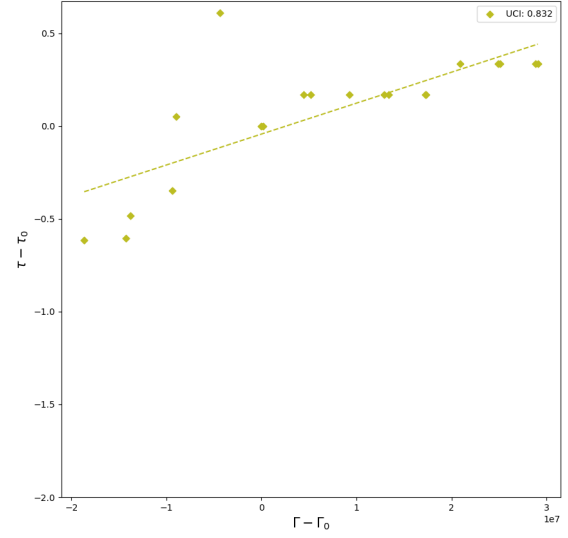
(a)



(b)



(c)



(d)

Figure A.59: Here are represented the important features about the phase transition in San Francisco for UCI = 0.832. In A.59a is represented the number of people trapped in the road network at time t . In A.59b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.59c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.597$. In A.59d we see the linear increase of τ compared to Γ

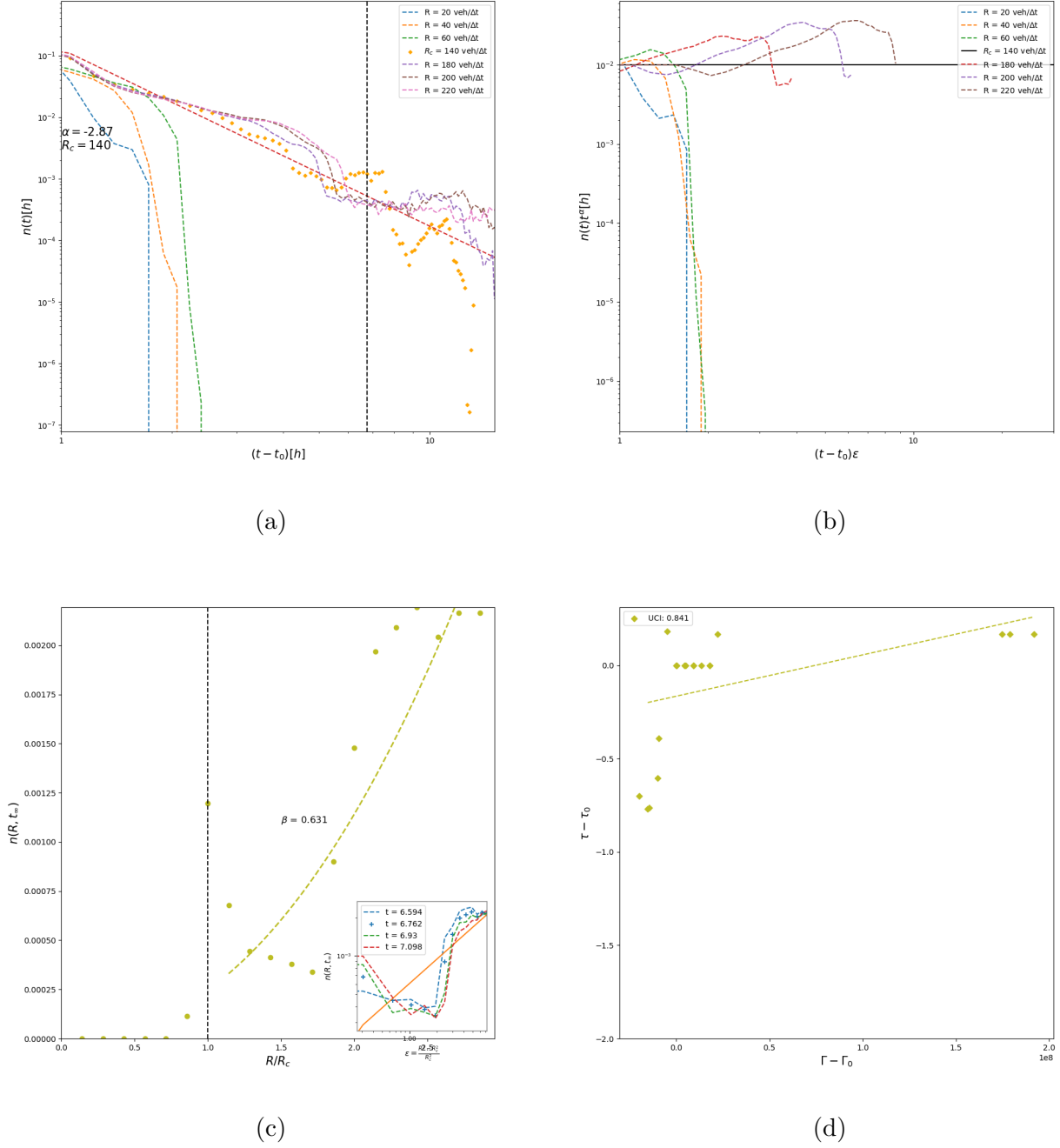


Figure A.60: Here are represented the important features about the phase transition in San Francisco for UCI = 0.841. In A.60a is represented the number of people trapped in the road network at time t . In A.60b, the vertical line representing t_∞ . we see the values $n(t)t^\alpha$ over $\epsilon(t - t_0)$, with $\epsilon = \frac{R^2 - R_c^2}{R_c^2}$ with t up to $t_{start} + t_{end_{fit}}$. In A.60c showing the phase transition requirement. We plot $n(R, t_\infty)$ with $\beta = 0.631$. In A.60d we see the linear increase of τ compared to Γ .

A.3.3 Population Distributions

Here a sketch of the distribution of populations and relative UCIs for the different cities at hand.

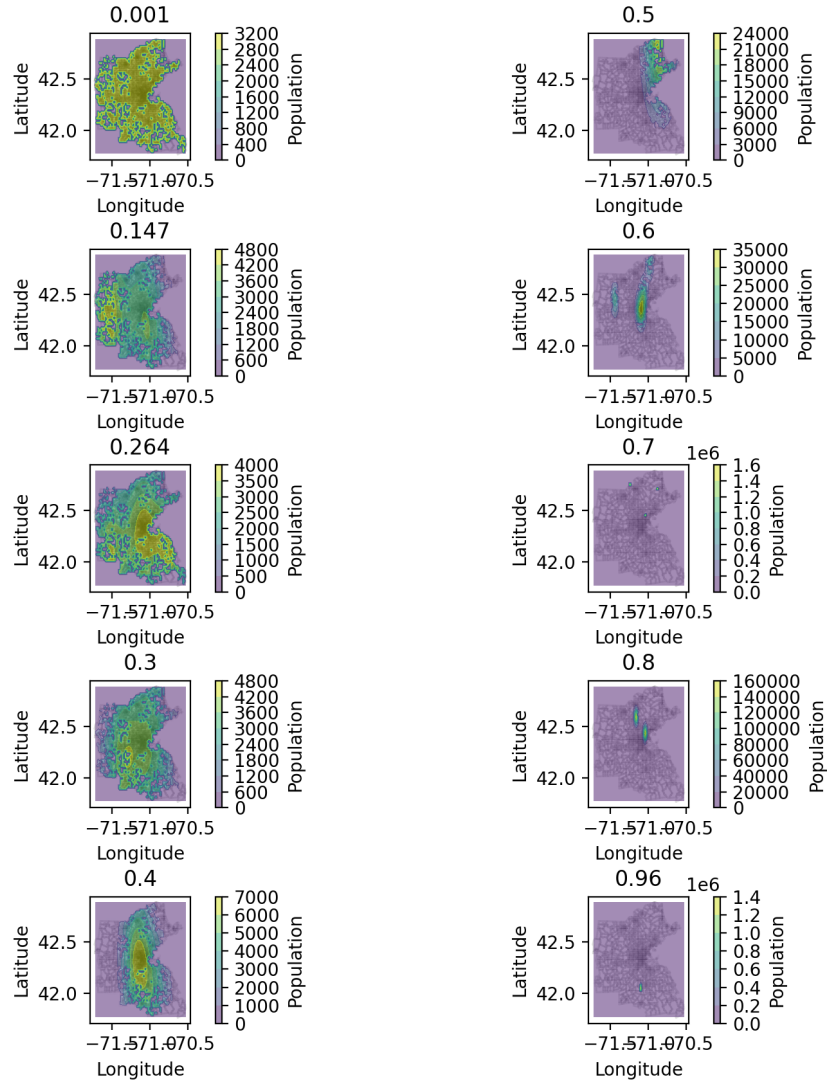


Figure A.61: Example of distribution of masses and their relative UCI for Boston. From right top to bottom left increasing monocentricity

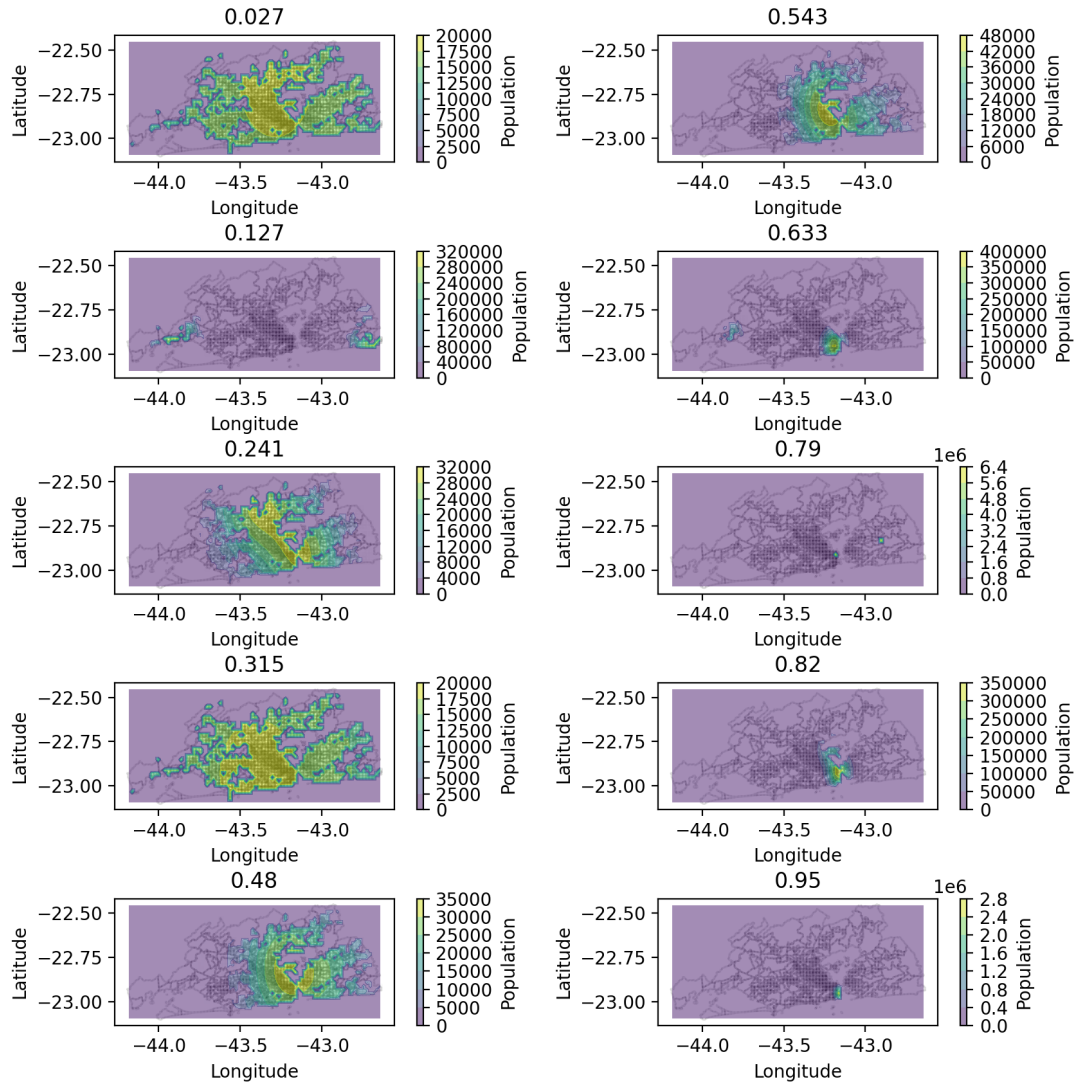


Figure A.62: Example of distribution of masses and their relative UCI for Rio de Janeiro. From right top to bottom left increasing monocentricity

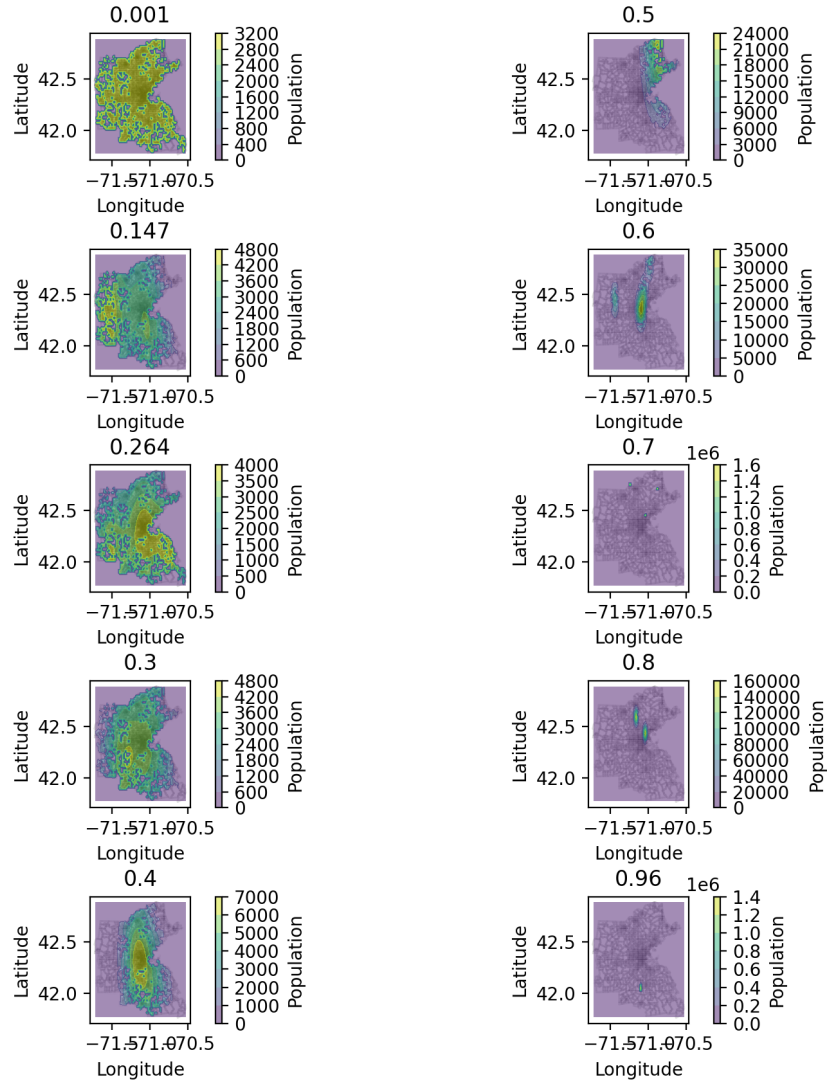


Figure A.63: Example of distribution of masses and their relative UCI for Los Angeles. From right top to bottom left increasing monocentricity

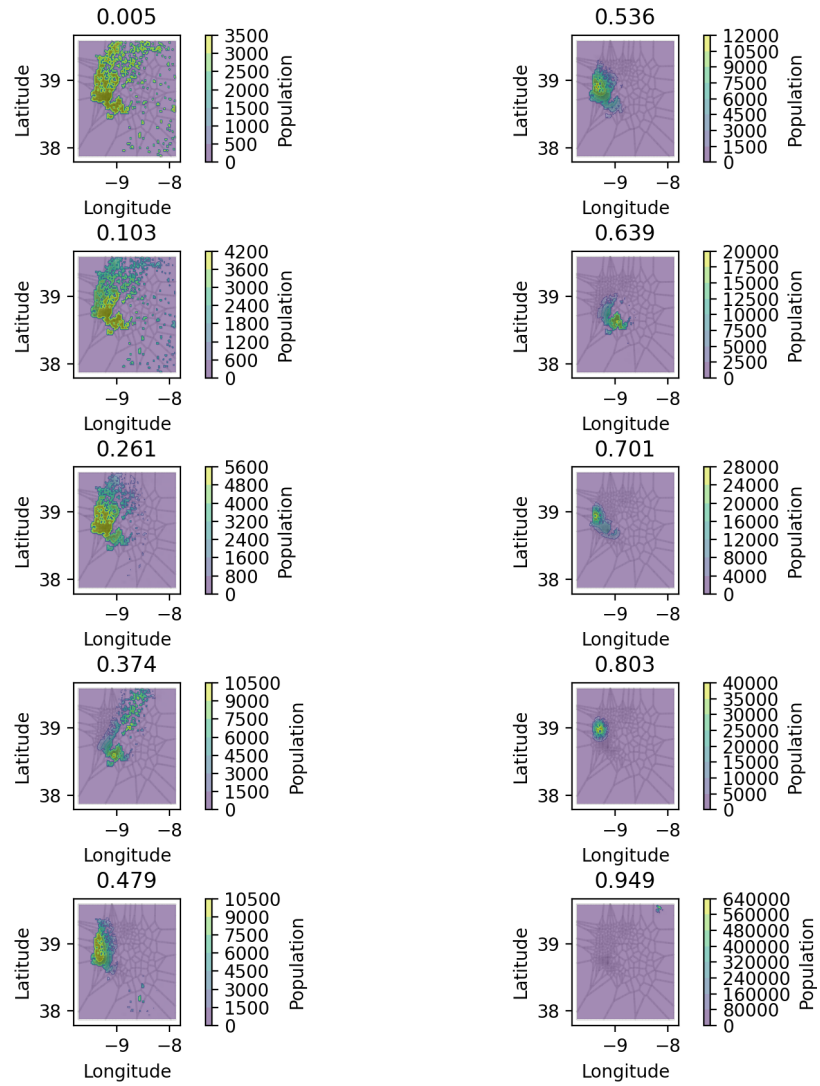


Figure A.64: Example of distribution of masses and their relative UCI for Lisbon. From right top to bottom left increasing monocentricity

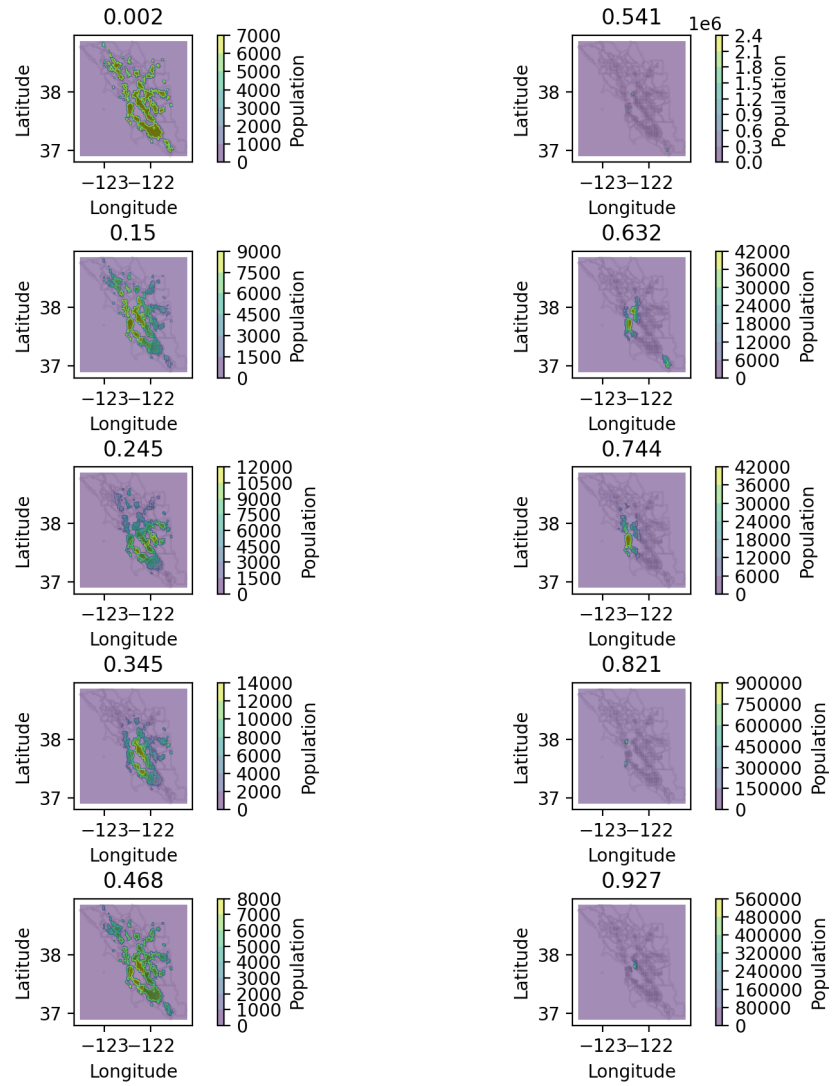


Figure A.65: Example of distribution of masses and their relative UCI for San Francisco. From right top to bottom left increasing monocentricity

A.4 Thanks

Last but not least, the very important task to thank people without which this work would neither have sense nor would have been realized. There are many things that I am glad for. First of all I have grown a lot and I am becoming more consistent and trustworthy in what I do thanks to each person I met during these years.

Prof. Bazzani, thanks for the patience, the guidance, share of resources and freedom you guaranteed me for these years.

Riccardo Gallotti, thanks for the guidance, patience, resources and freedom of these years. Thanks for having me in FBK.

Thanks Marta for the guidance, support, resources and freedom you gave me and having me. Thanks for the patience.

Thanks prof Rita Fioresi and Caligola project for the funding that enabled me to visit prof Gonzalez in Berkeley.

Thanks to Ing. Davide Micheli and Dott. Aldo Vannelli for sharing TIM data.

Thanks to Cuebic inc. for providing GPS data used in this study

Thanks to all the colleagues from Bologna to Trento to Berkeley.

Thanks to my family and my friends for the support. It has been a difficult trip that is finally ending.

Bibliography

- [1] <https://data.humdata.org/dataset/united-states-high-resolution-population-density-maps-demographic-estimates>.
- [2] <https://www.cdc.gov/museum/timeline/covid19.html>.
- [3] How will country-based mitigation measures influence the course of the covid-19 epidemic? *Lancet*, 395(10228):931–934, 2020 Mar 21. doi: doi:10.1016/S0140-6736(20)30567-5.
- [4] L. Alessandretti. What human mobility data tell us about COVID-19 spread. *Nature Reviews Physics*, 4(1):12–13, 2022. doi: 10.1038/s42254-021-00407-1.
- [5] L. Alessandretti, U. Aslak, and S. Lehmann. The scales of human mobility. *Nature*, 587: 402–407, 2020.
- [6] Alberto Aleta, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E Dean, M Elizabeth Halloran, Ira M Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro, and Yamir Moreno. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat. Hum. Behav.*, 2020. ISSN 2397-3374. doi: 10.1038/s41562-020-0931-9. URL <https://doi.org/10.1038/s41562-020-0931-9>.
- [7] L. Ambühl, M. Menendez, and M. C. González. Understanding congestion propagation by combining percolation theory with the macroscopic fundamental diagram. *Communications Physics*, 6(1):26, 2023. doi: 10.1038/s42005-023-01144-w. URL <https://doi.org/10.1038/s42005-023-01144-w>. Epub 2023 Feb 1.
- [8] Alex Arenas, Wesley Cota, Jesús Gómez-Gardeñes, Sergio Gómez, Clara Granell, Joan T Matamalas, David Soriano-Paños, and Benjamin Steinegger. Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Impact of Mobility and Social Distancing Interventions. *Phys. Rev. X*, 10(4):41055, dec 2020. doi: 10.1103/PhysRevX.10.041055. URL <https://link.aps.org/doi/10.1103/PhysRevX.10.041055>.
- [9] David Austin. Who’s number 1? hodge theory will tell us. <https://www.ams.org/publicoutreach/feature-column/fc-2012-12>, 2012.
- [10] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hud, José J. Ramasco, and Alessandro Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci. U. S. A.*, 106(51):21484–21489, 2009. ISSN 00278424. doi: 10.1073/pnas.0906910106.
- [11] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J. Ramasco, and Alessandro Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489, 2009. doi: 10.1073/pnas.0906910106. URL <https://www.pnas.org/doi/abs/10.1073/pnas.0906910106>.

- [12] Hugo Barbosa, Marc Barthélemy, Gourab Ghoshal, Charlotte R. James, Maxime Lenormand, Thomas Louail, Ronaldo Menezes, José J. Ramasco, Filippo Simini, and Marcello Tomasini. Human mobility: Models and applications. *Physics Reports*, 734:1–74, 2018. ISSN 0370-1573. doi: <https://doi.org/10.1016/j.physrep.2018.01.001>. URL <https://www.sciencedirect.com/science/article/pii/S037015731830022X>. Human mobility: Models and applications.
- [13] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge, UK, 2012.
- [14] Marc Barthélemy. *The Structure and Dynamics of Cities: Urban Data Analysis and Theoretical Modeling*. Cambridge University Press, 2016.
- [15] Marc Barthélemy. The role of parsimonious models in addressing mobility challenges. *npj Sustainable Mobility and Transport*, 1(1):11, 2024. ISSN 3004-8664. doi: 10.1038/s44333-024-00011-7. URL <https://doi.org/10.1038/s44333-024-00011-7>.
- [16] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3):1–101, 2011.
- [17] Marc Barthélemy, Pierre Bordin, Henri Berestycki, and Maurizio Griboaudi. Self-organization versus top-down planning in the evolution of a city. *Scientific Reports*, 3:2153, 2013.
- [18] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali. Smart cities of the future. *The European Physical Journal Special Topics*, 214(1):481–518, Nov 2012. ISSN 1951-6401. doi: 10.1140/epjst/e2012-01703-3. URL <https://doi.org/10.1140/epjst/e2012-01703-3>.
- [19] Armando Bazzani, Bruno Giorgini, Sandro Rambaldi, Riccardo Gallotti, and Luca Giovannini. Statistical laws in urban mobility from microscopic gps data in the area of florence. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(05):P05001, may 2010. doi: 10.1088/1742-5468/2010/05/P05001. URL <https://dx.doi.org/10.1088/1742-5468/2010/05/P05001>.
- [20] Enrico Bertuzzo, Lorenzo Mari, Damiano Pasetto, Stefano Miccoli, Renato Casagrandi, Marino Gatto, and Andrea Rinaldo. The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nat. Commun.*, 11(1):4264, 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-18050-2. URL <https://doi.org/10.1038/s41467-020-18050-2>.
- [21] Full W Bezdek JC, Ehrlich R. Fcm: the fuzzy c-means clustering algorithm. *Comput Geosci*, page 10(2–3):191, 1984.
- [22] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439(7075):462–465, January 2006. ISSN 1476-4687. doi: 10.1038/nature04292. URL <https://doi.org/10.1038/nature04292>.
- [23] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439(7075):462–465, 2006.
- [24] Mizzi C., Baroncini A., Fabbri A., Micheli D., Vannelli A., Criminisi C., Jean S., and Bazzani A. Individual mobility deep insight using mobile phones data. *EPJ Data Science*, 12:2193–1127, 2023. doi: 10.1140/epjds/s13688-023-00431-4. URL <https://doi.org/10.1140/epjds/s13688-023-00431-4>.

- [25] H. Barbosa F. Privitera R. Menezes R. Di Clemente C. Santana, F. Botta. Covid-19 is linked to changes in the time–space dimension of human mobility. *Nature Human Behavior*, 2023.
- [26] Francesco Calabrese, Mi Diao, Giovanni Di Lorenzo, Joseph Ferreira, and Carlo Ratti. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26:301–313, 2015.
- [27] Alessio Cardillo, Salvatore Scellato, Vito Latora, and Sergio Porta. Structural properties of planar graphs of urban street patterns. *Physical Review E*, 73(6):066107, 2006.
- [28] Serina Chang, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87, 2021. ISSN 1476-4687. doi: 10.1038/s41586-020-2923-3. URL <https://doi.org/10.1038/s41586-020-2923-3>.
- [29] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. Pastore Y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, and A. Vespignani. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. n.d. Manuscript.
- [30] Matteo Chinazzi, Jessica T. Davis, Ana Pastore y Piontti, Kunpeng Mu, Nicolò Gozzi, Marco Ajelli, Nicola Perra, and Alessandro Vespignani. A multiscale modeling framework for scenario modeling: Characterizing the heterogeneity of the covid-19 epidemic in the us. *Epidemics*, 47:100757, March 2024. doi: 10.1016/j.epidem.2024.100757. URL <https://doi.org/10.1016/j.epidem.2024.100757>.
- [31] D. Chowdhury and A. Schadschneider. Self-organization of traffic jams in cities: effects of stochastic dynamics and signal periods, 1998.
- [32] Marta Luisa Ciofi degli Atti, Stefano Merler, Caterina Rizzo, Marco Ajelli, Marco Mas-sari, Piero Manfredi, Cesare Furlanello, Gianpaolo Scalia Tomba, and Mimmo Iannelli. Mitigation Measures for Pandemic Influenza in Italy: An Individual Based Model Considering Different Scenarios. *PLoS One*, 3(3):e1790, mar 2008. URL <https://doi.org/10.1371/journal.pone.0001790>.
- [33] Thomas Courtat, Clément Gloaguen, and Stéphane Douady. Mathematics and morphogenesis of cities: A geometrical approach. *Physical Review E*, 83(3):036106, 2011.
- [34] Carlos F. Daganzo and Nikolas Geroliminis. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological*, 42(9):771–781, 2008. ISSN 0191-2615. doi: <https://doi.org/10.1016/j.trb.2008.06.008>. URL <https://www.sciencedirect.com/science/article/pii/S0191261508000799>.
- [35] P. Deville, C. Linard, S. Martin, et al. Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45):15888–15893, 2014.
- [36] Pablo Echenique, Jesús Gómez-Gardeñes, and Yamir Moreno. Dynamics of jamming transitions in complex networks. *arXiv preprint cond-mat/0412053*, Dec 2004. arXiv:cond-mat/0412053v1 [cond-mat.stat-mech].

- [37] Zhuangyuan Fan, Tianyu Su, Maoran Sun, Ariel Noyman, Fan Zhang, Alex ‘Sandy’ Pentland, and Esteban Moro. Diversity beyond density: Experienced social mixing of urban streets. *PNAS Nexus*, 2(4):pgad077, 03 2023. ISSN 2752-6542. doi: 10.1093/pnasnexus/pgad077. URL <https://doi.org/10.1093/pnasnexus/pgad077>.
- [38] Flavio Finger, Tina Genolet, Lorenzo Mari, Guillaume Constantin de Magny, Noël Magloire Manga, Andrea Rinaldo, and Enrico Bertuzzo. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. *Proc. Natl. Acad. Sci. U. S. A.*, 113(23):6421–6426, jun 2016. ISSN 1091-6490 (Electronic). doi: 10.1073/pnas.1522305113.
- [39] Bazzani A. Gallotti R. and Rambaldi S. Understanding the variability of daily travel-time expenditures using gps trajectory data. *EPJ Data Sci.*, 4(18), 2015. doi: <https://doi.org/10.1140/epjds/s13688-015-0055-z>.
- [40] M. T. Gastner and M. E.J. Newman. The spatial structure of networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 49(2):247–252, jan 2006. doi: 10.1140/epjb/e2006-00046-8. URL <https://doi.org/10.1140/2Fe2006-00046-8>.
- [41] Marino Gatto, Enrico Bertuzzo, Lorenzo Mari, Stefano Miccoli, Luca Carraro, Renato Casagrandi, and Andrea Rinaldo. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci.*, 117(19):10484 LP – 10491, may 2020. doi: 10.1073/pnas.2004978117. URL <http://www.pnas.org/content/117/19/10484.abstract>.
- [42] Nikolas Geroliminis and Carlos F. Daganzo. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770, 2008. URL <https://EconPapers.repec.org/RePEc:eee:transb:v:42:y:2008:i:9:p:759-770>.
- [43] N. M. Ferguson D. Laydon G. Nedjati-Gilani N. Imai K. Ainslie M. Baguelin S. Bhatia A. Boonyasiri Z. Cucunubá G. Cuomo-Dannenburg A. Dighe I. Dorigatti H. Fu K. Gaythorpe W. Green A. Hamlet W. Hinsley L. C. Okell S. Elstrand H. Thompson R. Verity E. Volz H. Wang Y. Wang P. G. T. Walker P. Winskill C. Whittaker C. Donnelly S. Riley A. C. Ghani. Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. *Report 9*, 2020.
- [44] Stefan Gheorghiu and Marc-Olivier Coppens. Heterogeneity explains features of ”anomalous” thermodynamics and statistics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(45):15852–15856, November 2004. ISSN 0027-8424. doi: 10.1073/pnas.0407191101. URL <https://europepmc.org/articles/PMC528778>.
- [45] F. Giannotti, M. Nanni, D. Pedreschi, et al. Unveiling the complexity of human mobility by querying and mining massive trajectory data. *VLDB Journal*, 20:695, 2011.
- [46] Genevieve Giuliano and Kenneth A. Small. Subcenters in the los angeles region. *Regional Science and Urban Economics*, 21(2):163–182, 1991. ISSN 0166-0462. doi: [https://doi.org/10.1016/0166-0462\(91\)90032-I](https://doi.org/10.1016/0166-0462(91)90032-I). URL <https://www.sciencedirect.com/science/article/pii/016604629190032I>.
- [47] M. C. Gonzalez, C. A. Hidalgo, and A. L. Barabasi. Understanding human mobility patterns. *Nature*, 454:779–782, 2008.

- [48] Marta C. Gonzalez, César A. Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *Nature*, 453:779–782, 2008.
- [49] P Grassberger, F Krause, and T von der Twer. A new type of kinetic critical phenomenon. *Journal of Physics A: Mathematical and General*, 17(3):L105, feb 1984. doi: 10.1088/0305-4470/17/3/003. URL <https://dx.doi.org/10.1088/0305-4470/17/3/003>.
- [50] Kappey J. Grenfell B. T., Bjørnstad O. N. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 2001.
- [51] D. Helbing. Traffic and related self-driven many-particle systems, 2000.
- [52] Dirk Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73(4):1067–1141, December 2001. ISSN 1539-0756. doi: 10.1103/revmodphys.73.1067. URL <http://dx.doi.org/10.1103/RevModPhys.73.1067>.
- [53] Dirk Helbing, editor. *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Understanding Complex Systems. Springer Berlin, Heidelberg, 1 edition, 2012. ISBN 978-3-642-24003-4. doi: 10.1007/978-3-642-24004-1. URL <https://doi.org/10.1007/978-3-642-24004-1>. Published: 06 May 2012 (Hardcover), 11 June 2014 (Softcover), 05 May 2012 (eBook).
- [54] ROBERT HERMAN, TENNY LAM, and ILYA PRIGOGINE. Kinetic theory of vehicular traffic: Comparison with data. *Transportation Science*, 6(4):440–452, 1972. ISSN 00411655, 15265447. URL <http://www.jstor.org/stable/25767676>.
- [55] Alberto Hernando, David Mateo, Jordi Bayer, and Ignacio Barrios. 2021. doi: 10.1101/2021.01.30.21250708.
- [56] Haye Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics*, 49(7):815–958, 2000. doi: 10.1080/00018730050198152. URL <https://doi.org/10.1080/00018730050198152>.
- [57] Emily Howerton, Lucie Contamin, Luke C. Mullany, Michelle Qin, Nicholas G. Reich, Samantha Bents, Rebecca K. Borchering, Sung mok Jung, Sara L. Loo, Claire P. Smith, John Levander, Jessica Kerr, J. Espino, Willem G. van Panhuis, Harry Hochheiser, Marta Galanti, Teresa Yamana, Sen Pei, Jeffrey Shaman, Kaitlin Rainwater-Lovett, Matt Kinsey, Kate Tallaksen, Shelby Wilson, Lauren Shin, Joseph C. Lemaitre, Joshua Kaminsky, Juan Dent Hulse, Elizabeth C. Lee, Clifton D. McKee, Alison Hill, Dean Karlen, Matteo Chinazzi, Jessica T. Davis, Kunpeng Mu, Xinyue Xiong, Ana Pastore y Piontti, Alessandro Vespignani, Erik T. Rosenstrom, Julie S. Ivy, Maria E. Mayorga, Julie L. Swann, Guido España, Sean Cavany, Sean Moore, Alex Perkins, Thomas Hladish, Alexander Pillai, Kok Ben Toh, Ira Longini, Shi Chen, Rajib Paul, Daniel Janies, Jean-Claude Thill, Anass Bouchnita, Kaiming Bi, Michael Lachmann, Spencer J. Fox, Lauren Ancel Meyers, Ajitesh Srivastava, Przemyslaw Porebski, Srinu Venkatramanan, Aniruddha Adiga, Bryan Lewis, Brian Klahn, Joseph Outten, Benjamin Hurt, Jiangzhuo Chen, Henning Mortveit, Amanda Wilson, Madhav Marathe, Stefan Hoops, Parantapa Bhattacharya, Dustin Machi, Betsy L. Cadwell, Jessica M. Healy, Rachel B. Slayton, Michael A. Johansson, Matthew Biggerstaff, Shaun Truelove, Michael C. Runge, Katriona Shea, Cécile Viboud, and Justin Lessler. Evaluation of the us covid-19 scenario modeling hub for informing pandemic response under uncertainty. *Nature Communications*, 14(1):7260, November 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-42680-x. URL <https://doi.org/10.1038/s41467-023-42680-x>.

- [58] <https://opendata.comune.bologna.it/explore/?sort=modified>.
- [59] [https://www.comune.bologna.it/novita/notizie/gemello digitale](https://www.comune.bologna.it/novita/notizie/gemello-digitale).
- [60] [https://www.comune.venezia.it/sites/comune.venezia.it/files/documenti/Turismo/Infografica_{nno2}](https://www.comune.venezia.it/sites/comune.venezia.it/files/documenti/Turismo/Infografica_nno2)
- [61] <https://www.fabbricadigitale.com/>.
- [62] <https://www.tim.it/>.
- [63] <https://www.venis.it/it>.
- [64] M. L. L. Iannini and Ronald Dickman. Kinetic theory of vehicular traffic. *American Journal of Physics*, 84(2):135–145, 02 2016. ISSN 0002-9505. doi: 10.1119/1.4935895. URL <https://doi.org/10.1119/1.4935895>.
- [65] Shan Jiang, Gaston A. Fiore, Yingxiang Yang, Joseph Ferreira, Emilio Frazzoli, and Marta C. González. A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing*, UrbComp ’13, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450323314. doi: 10.1145/2505821.2505828. URL <https://doi.org/10.1145/2505821.2505828>.
- [66] Raja Jurdak, Kun Zhao, Jiajun Liu, Maurice AbouJaoude, Mark Cameron, and David Newth. Understanding human mobility from twitter. *PLOS ONE*, 10(7):1–16, 07 2015. doi: 10.1371/journal.pone.0131469. URL <https://doi.org/10.1371/journal.pone.0131469>.
- [67] Shaw Darren J. Matthews Louise Chase-Topping Margo Haydon Dan T. Cornell Stephen J. Kappey Jens Wilesmith John Grenfell Bryan T. Keeling Matt J., Mark E. J. Woolhouse. Dynamics of the 2001 uk foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape. *American Association for the Advancement of Science*, 2001.
- [68] Boris S. Kerner. Three-phase traffic theory and highway capacity. *Physica A: Statistical Mechanics and its Applications*, 333:379–440, 2004. ISSN 0378-4371. doi: 10.1016/j.physa.2003.07.017. URL <https://www.sciencedirect.com/science/article/pii/S037843710300607X>. Available online at www.sciencedirect.com. Received 28 July 2003.
- [69] Arne Kesting, Martin Treiber, and Dirk Helbing. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1928):4585–4605, 2010. doi: 10.1098/rsta.2010.0084. URL <http://doi.org/10.1098/rsta.2010.0084>.
- [70] Moritz U G Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Open COVID-19 Data Working Group†, Louis du Plessis, Nuno R Faria, Ruoran Li, William P Hanage, John S Brownstein, Maylis Layan, Alessandro Vespignani, Huaiyu Tian, Christopher Dye, Oliver G Pybus, and Samuel V Scarpino. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science (80-.)*, 368(6490):493–497, may 2020. doi: 10.1126/science.abb4218. URL <https://doi.org/10.1126/science.abb4218>.
- [71] Daniel Krajzewicz. Traffic simulation with sumo - simulation of urban mobility. In *Fundamentals of Traffic Simulation*, International Series in Operations Research and Management Science, pages 269–294. Springer, 2010. ISBN 978-1-4419-6141-9.

- [72] Robert Kölbl and Dirk Helbing. Energy laws in human travel behaviour. *New Journal of Physics*, 5(1):48, may 2003. doi: 10.1088/1367-2630/5/1/348. URL <https://dx.doi.org/10.1088/1367-2630/5/1/348>.
- [73] P. Coletti N. Hens V. Colizza L. D. Domenico, G. Pullano. Expected impact of school closure and telework to mitigate covid-19 epidemic in france. page 1–15, 2020.
- [74] Daqing Li, Bo Fu, Ying Wang, Guanrong Lu, Yevgeny Berezin, H. Eugene Stanley, and Shlomo Havlin. Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences of the United States of America*, 112(3):669–672, 2015. doi: 10.1073/pnas.1419185112. URL <https://doi.org/10.1073/pnas.1419185112>. Epub 2014 Dec 31.
- [75] Whitham G. B. Lighthill M. J. kinematic waves ii. a theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond.*, 1955.
- [76] Erjian Liu, Mattia Mazzoli, Xiao-Yong Yan, and José J. Ramasco. A generalized vector-field framework for mobility. *Communications Physics*, 7(1):190, 2024. doi: 10.1038/s42005-024-01672-z. URL <https://doi.org/10.1038/s42005-024-01672-z>.
- [77] Thomas Louail, Maxime Lenormand, Oliva G. Cantu Ros, Miguel Picornell, Ricardo Herranz, Enrique Frias-Martinez, José J. Ramasco, and Marc Barthelemy. From mobile phone data to the spatial structure of cities. *Scientific Reports*, 4(1), June 2014. ISSN 2045-2322. doi: 10.1038/srep05276. URL <http://dx.doi.org/10.1038/srep05276>.
- [78] Thomas Louail, Maxime Lenormand, Olga Garcia Cantu Ros, Miquel Picornell, Ricardo Herranz, Enrique Frias-Martinez, Jose J Ramasco, and Marc Barthélemy. Uncovering the spatial structure of mobility networks. *Nature Communications*, 6:6007, 2015.
- [79] Rémi Louf and Marc Barthelemy. Modeling the polycentric transition of cities. *Physical review letters*, 111:198702, 11 2013. doi: 10.1103/PhysRevLett.111.198702.
- [80] Rémi Louf and Marc Barthélemy. A typology of street patterns. *Journal of The Royal Society Interface*, 11(101):20140924, 2014.
- [81] L. Lucchini, S. Centellegher, L. Pappalardo, R. Gallotti, F. Privitera, B. Lepri, and M. de Nadai. Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures. *Scientific Reports*.
- [82] Stefan Lämmer, Björn Gehlsen, and Dirk Helbing. Scaling laws in the spatial structure of urban road networks. *Physica A: Statistical Mechanics and its Applications*, 363(1):89–95, 2006.
- [83] Luis E. Olmos Serdar Çolak Sajjad Shafiei Meead Saberi Marta C. González. Macroscopic dynamics and the collapse of urban traffic. *PNAS*, 2023.
- [84] AP Masucci, Darren Smith, Andrew Crooks, and Michael Batty. Random planar graphs and the london street network. *The European Physical Journal B*, 71(2):259–271, 2009.
- [85] M Mazzoli, A Molas, A Bassolas, et al. Field theory for recurrent mobility. *Nature Communications*, 10:3895, 2019. doi: 10.1038/s41467-019-11841-2.
- [86] Simone Meinardi, Paul Nissenson, Barbara Barletta, Donald Dabdub, F. Sherwood Rowland, and Donald R. Blake. Influence of the public transportation system on the air quality of a major urban center. a case study: Milan, italy. *Atmospheric Environment*, 42(34):7915–7923, 2008. ISSN 1352-2310. doi: <https://doi.org/10.1016/j.atmosenv.2008.07.046>. URL <https://www.sciencedirect.com/science/article/pii/S1352231008006341>.

- [87] Diamanti R Micheli D. Statistical analysis of interference in a real lte access network by massive collection of mdt radio measurement data from smartphones. *PhotonIcs Electromagnetics Research Symposium—Spring(PIERS-Spring)*, page 1906, 2019.
- [88] Vannelli A Micheli D, Muratore G. Rain effect on 4g lte in-car electromagnetic propagation analyzed through mdt radio data measurement reported by mobile phones. *IEEE Trans Antennas Propag*, page 69(12):8641, 2021.
- [89] C Mizzi, A Fabbri, G Colombini, F Bertini, and A Bazzani. A survival model to explain the statistical properties of multimodal mobility. *Journal of Statistical Mechanics: Theory and Experiment*, 2022(2):023404, feb 2022. doi: 10.1088/1742-5468/ac4c40. URL <https://dx.doi.org/10.1088/1742-5468/ac4c40>.
- [90] Chiara Mizzi. Big data analytics and modeling for human mobility. 2023.
- [91] Chiara Mizzi, Fabbri Alessandro, Sandro Rambaldi, and et al. Unraveling pedestrian mobility on a road network using icts data during great tourist events. *EPJ Data Science*, 7(44): 21, 2018. doi: 10.1140/epjds/s13688-018-0168-2. URL <https://doi.org/10.1140/epjds/s13688-018-0168-2>.
- [92] Chiara Mizzi, Alex Baroncini, Alessandro Fabbri, Davide Micheli, Aldo Vannelli, Carmen Criminisi, Susanna Jean, and Armando Bazzani. Individual mobility deep insight using mobile phones data. *EPJ Data Science*, 12(1):56, 2023. doi: 10.1140/epjds/s13688-023-00431-4. URL <https://doi.org/10.1140/epjds/s13688-023-00431-4>.
- [93] Esteban Moro, Dan Calacci, Xiaowen Dong, and Alex Pentland. Mobility patterns are associated with experienced income segregation in large us cities. *Nature Communications*, 12(1):4633, July 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24899-8. URL <https://doi.org/10.1038/s41467-021-24899-8>.
- [94] Fionn Murtagh and Pierre Legendre. Ward’s hierarchical agglomerative clustering method: Which algorithms implement ward’s criterion? *Journal of Classification*, 31(3):274–295, October 2014. ISSN 1432-1343. doi: 10.1007/s00357-014-9161-z. URL <http://dx.doi.org/10.1007/s00357-014-9161-z>.
- [95] K. Nagel and Schreckenberg M. A cellular automaton model for freeway traffic. *J. Phys. I France*, (2):2221–2229, 1992. doi: <https://doi.org/10.1051/jp1:1992277>.
- [96] P. Nouvellet et al. Reduction in mobility and COVID-19 transmission. *Nat. Commun.*, 12: 1090, 2021.
- [97] Eiman Nuaimi, Hind Neyadi, Nader Mohamed, and Jameela Al-Jaroodi. Applications of big data to smart cities. *Journal of Internet Services and Applications*, 6, 08 2015. doi: 10.1186/s13174-015-0041-5.
- [98] Nuria Oliver, Bruno Lepri, Harald Sterly, Renaud Lambiotte, Sébastien Deletaille, Marco De Nadai, Emmanuel Letouzé, Albert Ali Salah, Richard Benjamins, Ciro Cattuto, Vittoria Colizza, Nicolas de Cordes, Samuel P. Fraiberger, Till Koebe, Sune Lehmann, Juan Murillo, Alex Pentland, Phuong N Pham, Frédéric Pivetta, Jari Saramäki, Samuel V. Scarpino, Michele Tizzoni, Stefaan Verhulst, and Patrick Vinck. Mobile phone data for informing public health actions across the covid-19 pandemic life cycle. *Science Advances*, 6(23):eabc0764, 2020. doi: 10.1126/sciadv.abc0764. URL <https://www.science.org/doi/abs/10.1126/sciadv.abc0764>.

- [99] Luis E Olmos, Serdar Çolak, Saman Shafiei, Meead Saberi, and Marta C González. Macroscopic dynamics and the collapse of urban traffic. *Proceedings of the National Academy of Sciences*, 115(50):12654–12661, 2018.
- [100] del Rio-Chanona R. M. Pichler A. Martín-Corral D. Chinazzi M. Lafond F. Ajelli M. Moro E. Moreno Y. Vespignani A. Farmer J. Doyne Pangallo M., Aleta A. The unequal effects of the health–economy trade-off during the covid-19 pandemic. *Nature Human Behaviour*, 2024.
- [101] Luca Pappalardo, Filippo Simini, Salvatore Rinzivillo, Dino Pedreschi, Fosca Giannotti, and Albert-Laszlo´ Barabasi. Returners and explorers dichotomy in human mobility.
- [102] R.H.M. Pereira, V. Nadalin, L. Monasterio, and P.H.M. Albuquerque. Urban centrality: A simple index. *Geographical Analysis*, 45:77–89, 2013. doi: 10.1111/gean.12002.
- [103] Gallotti R, Bazzani A, Rambaldi S, and Barthelemy M. A stochastic model of randomly accelerated walkers for human mobility. *Nature Communications*, 2016.
- [104] Injong Rhee, Minsu Shin, Seongik Hong, Kyunghan Lee, Seong Joon Kim, and Song Chong. On the levy-walk nature of human mobility. *IEEE/ACM Transactions on Networking*, 19(3): 630–643, 2011. doi: 10.1109/TNET.2011.2120618.
- [105] Micheli D Scalon A. Estimation of mobility direction of a people flux by using a live 3g radio access network and smartphones in non-connected mode. *Proc. IEEE 15th Int. Conf. Environ. Elect. Eng.*, page 1869, 2015.
- [106] Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, and Cecilia Mascolo. Socio-spatial properties of online location-based social networks. page 329 – 336, 2011. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124442007&partnerID=40&md5=fa8cacd28116c1e2996bb4f4625bc3b0>. Cited by: 348.
- [107] Frank Schlosser, Benjamin F Maier, Olivia Jack, David Hinrichs, Adrian Zachariae, and Dirk Brockmann. COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proc. Natl. Acad. Sci.*, 117(52):32883–32890, dec 2020. doi: 10.1073/pnas.2012326117. URL <https://doi.org/10.1073/pnas.2012326117>.
- [108] Michael Schreckenberg Shanshan Wang and Thomas Guhr. Identifying subdominant collective effects in a large motorway network. *Journal of Statistical Mechanics: Theory and Experiment*, 2022(11):113402, 2022.
- [109] Michael Schreckenberg-Thomas Guhr Shanshan Wang, Sebastian Gartzke. Quasi-stationary states in temporal correlations for traffic systems: Cologne orbital motorway as an example. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(10):103404, 2020.
- [110] Filippo Simini, Marta C González, Amos Maritan, and Albert-László Barabási. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, 2012.
- [111] Filippo Simini, Gianni Barlacchi, Massimiliano Luca, and Luca Pappalardo. A Deep Gravity model for mobility flows generation. *Nature Communications*, 12(1):6576, 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-26752-4. URL <https://doi.org/10.1038/s41467-021-26752-4>.
- [112] Chaoming Song, Tal Koren, PuWang, and Albert-László Barabási. Modelling the scaling properties of human mobility. *Nature Physics*, 6:818–823, 2010.

- [113] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. Limits of predictability in human mobility. *Science*, 327:1018–1021, 2010.
- [114] Wilbur R. Thompson. The coefficient of localization: An appraisal. *Southern Economic Journal*, 23(3):320–325, 1957. ISSN 00384038. URL <http://www.jstor.org/stable/1054222>.
- [115] Michele Tizzoni, Paolo Bajardi, Adeline Decuyper, Guillaume Kon Kam King, Christian M. Schneider, Vincent Blondel, Zbigniew Smoreda, Marta C. González, and Vittoria Colizza. On the use of human mobility proxies for modeling epidemics. *PLOS Computational Biology*, 10(7): 1–15, 07 2014. doi: 10.1371/journal.pcbi.1003716. URL <https://doi.org/10.1371/journal.pcbi.1003716>.
- [116] Alessandro Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nature Physics*, 8(1):32–39, Jan 2012. ISSN 1745-2481. doi: 10.1038/nphys2160. URL <https://doi.org/10.1038/nphys2160>.
- [117] Feilong Wang, Daqing Li, Xiaoyun Xu, Ruoqian Wu, and Shlomo Havlin. Percolation properties in a traffic model. *Europhysics Letters*, 112(3):033401, 2021. doi: 10.1088/1742-5468/abdf5e. URL <https://doi.org/10.1088/1742-5468/abdf5e>.
- [118] Shenhao Wang, Yunhan Zheng, Guang Wang, Takahiro Yabe, Esteban Moro, and Alex ‘Sandy’ Pentland. Infrequent activities predict economic outcomes in major american cities. *Nature Cities*, 1(4):305–314, April 2024. ISSN 2731-9997. doi: 10.1038/s44284-024-00051-7. URL <https://doi.org/10.1038/s44284-024-00051-7>.
- [119] Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow, and Caroline O. Buckee. Quantifying the impact of human mobility on malaria. *Science*, 338(6104):267–270, October 2012. ISSN 0036-8075. doi: 10.1126/science.1223467.
- [120] Yanyan Xu, Luis E. Olmos, David Mateo, Alberto Hernando, Xiaokang Yang, and Marta C. Gonzalez. Urban dynamics through the lens of human mobility. 2023.
- [121] James Demmelb-Samuel Williams Xuan Jianga, Raja Senguptaa. Lpsim: Large scale multi-gpu parallel computing based regional scale traffic simulation framework.
- [122] Takahiro Yabe, Massimiliano Luca, Kota Tsubouchi, Bruno Lepri, Marta C. Gonzalez, and Esteban Moro. Enhancing human mobility research with open and standardized datasets. *Nature Computational Science*, 4(7):469–472, July 2024. ISSN 2662-8457. doi: 10.1038/s43588-024-00650-3. URL <https://doi.org/10.1038/s43588-024-00650-3>.
- [123] Chuantao Yin, Zhang Xiong, Hui Chen, Jingyuan Wang, Daven Cooper, and Bertrand David. A literature survey on smart cities. *Science China Information Sciences*, 58, 08 2015. doi: 10.1007/s11432-015-5397-4.
- [124] J. Zhang, M. Litvinova, Y. Liang, W. Zheng, H. Shi, A. Vespignani, C. Viboud, M. Ajelli, and H. Yu. The impact of relaxing interventions on human contact patterns and SARS-CoV-2 transmission in china. *Sci Adv*, 7(19):eabe2584, 2021. doi: 10.1126/sciadv.abe2584.
- [125] C. Zhao, A. Zeng, and C. H. Yeung. Characteristics of human mobility patterns revealed by high-frequency cell-phone position data. *EPJ Data Science*, 10:5, 2021.
- [126] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, 23(3):337–343, 1977. doi: 10.1109/TIT.1977.1055714.