University of Trento
Department of Mathematics
Ph.D. in Mathematics

XXXIV Doctoral Cycle

Network models for large-scale human mobility

Advisors
Prof. Manlio De Domenico
Prof. Claudio Agostinelli

Ph.D. Candidate
Sebastian Raimondo

Academic Year 2021-2022

Abstract

Human mobility is a complex phenomenon emerging from the nexus between social, demographic, economic, political and environmental systems. In this thesis we develop novel mathematical models for the study of complex systems, to improve our understanding of mobility patterns and enhance our ability to predict local and global flows for real-world applications.

The first and second chapters introduce the concept of human mobility from the point of view of complex systems science, showing the relation between human movements and their predominant drivers. In the second chapter in particular, we will illustrate the state of the art and a summary of our scientific contributions. The rest of the thesis is divided into three parts: *structure*, *causes* and *effects*.

The third chapter is about the *structure* of a complex system: it represents our methodological contribution to Network Science, and in particular to the problem of network reconstruction and topological analysis. We propose a novel methodological framework for the definition of the topological descriptors of a complex network, when the underlying structure is uncertain. The most used topological descriptors are redefined – even at the level of a single node – as probability distributions, thus eluding the reconstruction phase. With this work we have provided a new approach to study the topological characteristics of complex networks from a probabilistic perspective.

The forth chapter deals with the *effects* of human mobility: it represents our scientific contribution to the debate about the COVID-19 pandemic and its consequences. We present a complex-causal analysis to investigate the relationship between environmental conditions and human activity, considered as the components of a complex socio-environmental system. In particular, we derive the network of relations between different flavors of human mobility data and other social and environmental variables. Moreover, we studied the effects of the restrictions imposed on human mobility – and human activities in general – on the environmental system. Our results highlight a statistically significant qualitative improvement in the environmental variable of interest, but this improvement was not caused solely by the restrictions due to COVID-19 pandemic, such as the lockdown.

The fifth and sixth chapters deal with the modelling of *causes* of human mo-

bility: the former is a concise chapter that illustrate the phenomenon of human displacements caused by environmental disasters. Specifically, we analysed data from different sources to understand the factors involved in shaping mobility patterns after tropical cyclones. The latter presents the Feature-Enriched Radiation Model (FERM), our generalization of the Radiation Model which is a state-of-the-art mathematical model for human mobility. While the original Radiation Model considers only the population as a proxy for mobility drivers, the FERM can handle any type of exogenous information that is used to define the attractiveness of different geographical locations. The model exploits this information to divert the mobility flows towards the most attractive locations, balancing the role of the population distribution. The mobility patterns at different scales can be reshaped, following the exogenous drivers encoded in the features, without neglecting the global configuration of the system.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

> Pensa costantemente all'universo come a una creatura vivente unica, che racchiude una sola sostanza e una sola anima; pensa come tutto sia assorbito in una sola sensazione di questa creatura; come tutto si compia grazie a un unico impulso, e tutte le cose siano causa comune di cio che nasce, e quali siano il loro concatenamento e la loro connessione.
>
> *A sé stesso*, Marco Aurelio

Life on Earth is a story of the interplay and adaptation to the most varied conditions. The ability to move in the environment increases the opportunities of the organisms to grow and flourish. Human beings have developed increasingly sophisticated mobility systems that allow them to search for optimal living conditions and to propagate ideas, innovations, but also infectious diseases and conflicts [1]. In the current context of globalization, human migration is increasingly influenced by the broader global transformations in economy, society and environment. Concurrently, migration has its own important effects on shaping these transformations. Therefore, the scientific understanding of human migration has significant implications on both theoretical and practical fronts. On the one hand, the theory of human mobility and migration plays an important role in the study of human behaviour, epidemiology, demography, economy and social science [2, 3, 4, 5, 6]. On the other hand, the description and prediction of large-scale human movements is necessary to implement correct policies to absorb the impacts that migrants have on the locations of origin and destination (consider for example the effect on labor markets [7], or the effect of mobility on epidemic

spreading [8]) and also to ensure a safe, orderly and humane migration to the migrants [9]. Hence, in this work, we aim to improve our understanding of the complex systems that give rise to mobility patterns and to enhance our ability to predict local and global flows to facilitate pro-active policy planning beneficial to migrants and local populations.

Human mobility is a complex phenomenon in both its causes and effects. Usually, social and economical drivers are considered as the main causes of large-scale mobility, but also political, environmental and demographic conditions are crucial causal factors. Moreover, the relation and mutual influence between these drivers is crucial in determining the migrants choice to move [10]. Accordingly, large-scale human mobility patterns originate from micro-scale individual choices, which in turn are the consequences of a complex relation between different drivers at different scales. We will describe the mobility patterns looking at the macro-scale behaviour of the whole complex system, assuming – in the words of Anderson [11] – the "extensive research" perspective, proper of complex system science.

Thesis structure – In the first and second chapters of this thesis we will introduce the concept of human mobility from the point of view of complex systems science, showing in more detail the relation between human movements and their predominant drivers. In the second chapter in particular, we will illustrate the main mathematical models used to describe human mobility and causal structures, and we will conclude presenting our contributions. The rest of the thesis can be ideally divided in three parts: *structure, causes* and *effects*.

The third chapter is about the *structure* of a complex system: it represents our methodological contribution to the network science, and in particular to the problem of network reconstruction and topological analysis. We propose a novel methodological framework to evaluate the uncertainty related to the structural features of complex systems, replacing the topological descriptors of the underlying network – even at the level of a single node – with appropriate probability distributions, eluding the reconstruction phase, up to now indispensable. Our results provide a grounded framework for the analysis and the interpretation of widely used topological descriptors, such as the degree centrality, clustering and clusters, in scenarios where the existence of network connectivity is statistically inferred or when the probabilities of existence π_{ij} of the edges are known. To this purpose we also provide a simple and mathematically grounded procedure to transform the discriminating statistics into the probabilities π_{ij}. This chapter corresponds to the paper *"Measuring topological descriptors of complex networks under uncertainty"*, published in *Physical Review E* in 2021 [12].

The forth chapter deals with the *effects* of human mobility: the period of my Ph.D. was signed by the COVID-19 pandemic, during which a great part of

the scientific community has been called to support institutions and inform citizens in order to control epidemic spreading and also to analyse and synthesise observations from a wide spectrum of related phenomena: changes in mobility, environment and economy, the waves of infodemics, the rising problems in social-psychology and political sciences, among the others [13, 14, 15, 16, 17]. Under these circumstances, we urged to give our scientific contribution to the community leveraging on the mathematical methods mastered during the first year of Ph.D. Therefore, in the third chapter we present a complex-causal analysis to investigate the relationship between environmental conditions and human activity – with human mobility at the first place – considered as the components of a complex socio-environmental system. We apply information theory, network science and Bayesian inference to analyze their structural relations and nonlinear dynamics between January 2019 and August 2020 in Northern Italy, i.e., before, during and after the national lockdown. We unravel the causal relationships between the 16 environmental conditions and human activity variables, mapping the backbone of the complex interplay between the physical observables to a causal network model. To identify a tipping point during the period of observation, denoting the presence of a regime shift between distinct network states (i.e., before and during the shock) we introduce a novel information-theoretic method based on statistical divergence widely used in statistical physics. We find that despite a measurable decrease in NO2 concentration, due to an overall decrease in human activities, targeted mobility restrictions and the reduction of social activities are inappropriate remedies to reduce emissions. Our results provide a functional characterization of socio-environmental interdependent systems and our analytical framework can be used, more generally, to characterize environmental changes and their interdependencies using statistical physics. This chapter corresponds to the paper *"COVID-19 lockdown unravels the complex interplay between environmental conditions and human activity"*, published in *Complexity* in 2022 [18].

The fifth and sixth chapters deal with the modelling of *causes* of human mobility: the former is a more concise chapter that consists of an illustrative example of large scale mobility caused by environmental disasters. The latter deals with the generalization of the Radiation Model, a mathematical model for human mobility. The causal process that leads to the emergence of complex mobility patterns, can be modeled as an emission-absorption process that takes place between radiative bodies representing space locations. A well known mathematical model that formalizes this process is the "radiation model" [19]. The radiation model is a parameter free model used to predict mobility flows where no other data except population is available. The flows are determined just on the basis of the spatial distribution of population, that is a proxy for all other causal drivers, and no other exogenous information can be included. In the sixth chapter we present

our generalization that explicitly include external drivers of migration represented by the features of the nodes of mobility networks. Our Feature-Enriched Radiation Model (FERM) maintains the same physical process of the original model, but generalizes its stochastic process in order to drive the flows according to the exogenous information. We describe the mathematical formulation of our model and the numerical method for the computation of the mobility network structure, along with some numerical experiments. The content of the sixth chapter was presented as a contributed talk in Lyon, at the *Conference on Complex Systems – CCS2021*, and the corresponfing paper is to be submitted.

The concluding chapter summarizes the main results of the thesis and speculate on possible future research directions.

1.1 The emergence of human mobility

> Se ciò non producesse nulla sull'anima, anche questo piccolo incremento non produrrebbe nulla e la totalità neppure.
>
> G. W. Leibniz, Scritti filosofici

What is a complex system? Attempts to answer this question are frequently found in the introduction of scientific papers and thesis, if not the main topic of entire books [20, 21, 22, 23]. The definition of what a complex system is, still represents a hot topic for scientists and philosophers [24, 25, 26]. During the past century, different conceptions of complex systems have been proposed by mathematicians, physicists, engineers, biologists, social scientists, but the last word on a unifying framework is still missing. Nonetheless, complexity scientists are aware of a set of properties that a system should have in order to be defined *complex*: numerous agents, network of relations between agents, non-linear dynamic, hierarchical organization and adaptive behaviour are some of the best known characteristics. In addition to these, one of the most distinctive property that a complex system posses is *emergence*. One might say that emergence is the reason why complex systems are appealing. The definition of emergence in it self is obscure and still largely debated [27, 28, 29]. Arguably, the first intuition of "emergence" was conceived by Aristotle in his *Metaphysics*: "In the case of all things which have several parts and in which the totality is not, as it were, a mere heap, but *the whole is something beside the parts*, there is a cause". Something besides the parts *emerges* because the totality is not a mere heap, but the way in which the elements interact to become a whole plays a fundamental role. On a strictly similar concept are

based the theoretical foundations of the Gestalt psychology, a school of psychology risen in the early twentieth century in Austria and Germany, that proposes the "gestalt" as a new category. Kurt Koffka, one of the leading proponent of the theory, argues that "to apply the category of cause and effect means to find out which parts of nature stand in this relation. Similarly, to apply the gestalt category means to find out which parts of nature belong as parts to functional wholes, to discover their position in these wholes, their degree of relative independence, and the articulation of larger wholes into sub-wholes" [30]. A more explicit definition of emergence, that may be useful to better frame emergence in our context, is given by Peter Corning in [28]: he proposed that emergent phenomena are "synergistic effects" generated by cooperative interactions between "functional complementarities". From those interactions a "qualitative novelty" arise, so that the components of the system are modified, reshaped, or transformed by their participation in the whole. Thus, an emergent property is a qualitative novelty of the system and as such it is not identical with, reducible to, or deducible from the parts that compose the system, but is a characteristic that is conceivable only by considering the system as a whole. Different studies gave ground to a more formal definition of emergence which may help in establish a novel methodology for the study of complex systems (see e.g. [31, 32, 33, 34]). Emergence is in fact a disruptive concept that may have a crucial impact on the modern scientific worldview. It is a pivotal concept in building a potential alternative to reductionism, that has been a basic foundation of the scientific methodology since the XVII century (see e.g. [35, 29, 36] for an overview). Examples of emergent behaviour are the emergence of harmony from the combination of musical notes, the emergence of collective behaviour in honeybee clusters [37], flocking birds [38] and school fish [39], the emergence of life from pre-biotic components [40, 41] the emergence of consensus in social systems from the interaction of people without any central authority [42].

The study of emergent phenomena give rise to still debated questions on causality [43], that stand at the base of the (network) structure of complex systems [44]. The above mentioned "cooperative interactions" between the components of a complex system, originate from causal relations of different kind [45, 34], that connect the components in a network structure. Each component is a node of the network, and the relations between them are represented by the edges. The collective behaviour of a complex system and its emergent phenomena are determined by the complex interplay between the dynamics of its constituent parts and the network structure of the relation between them [46]. Therefore, causality (and causal inference) plays a central role in defining emergent phenomena (see e.g. the concept of downward causation in [31]) and in general in the study of the interactions between parts of a complex system [47].

Within complex system science, human mobility can be considered as a phe-

nomenon emerging from the complex nexus between social, demographic, economic, political and environmental systems [48, 10]. The actual conditions of these systems in a certain area determine (causally) the opportunities and the costs for individuals to move in other places. The current global and local transformations in society and environment portend an unprecedented uncertainty in the evolution of large scale mobility and migration [49]. At the date of this thesis, the number of international migrants was estimated to be almost 281 million globally in 2020, 9 million more than in 2019, confirming the unceasing growth of the last decades. The COVID-19 pandemic is not halting migration, but just slowing down the pace [50]. At the contrary, the effect of political instability, growing inequality and environmental degradation are triggering large-scale displacements in vulnerable countries [49]. For these reasons, some scholars are debating whether large-scale human mobility and migration is becoming an *emergency*, besides being an *emergence* (see e.g. [51, 52]). Moreover, pundits are struggling to define what is a migrant, what is a refugee, what are the differences between an economic migrant, an environmental migrant or a climate refugee. These definitions would have a distinct meaning only if associated with a precise cause of migration. But as said before, human mobility emerges from the interplay between innumerable, interdependent variables. A better definition of the causes of migration, and how they relate to complex systems is the argument of the coming section.

1.2 Complexity and causality in human mobility

What is the cause of human mobility and migration? On the micro-scale, mobility and migration are generally caused by the individuals search for satisfaction of some need. Better economic conditions, social ties, leaving a dangerous area plagued by conflicts or natural disasters are just a few examples. These individual choices are reflected in the patterns observed at larger scales. Several scientific theories of migration have been proposed since the end of the XIX century to explain mass human movements. The foundational papers "The Laws of Migration" by E. G. Ravenstein in 1885 and 1889 [53, 54] laid the first stone of the scientific discussion on migration processes. In these works, the author gave some general "laws of migration" based on geographical, social and economical principles. In particular, the laws highlight the effect of distance and the influence of the economic factors, that are considered as the main causes of migration. Similarly, subsequent economic theories of migration, stemming from neo-classic economics, suggest that migration decisions are consequences of rational choices of individuals aimed at maximize utility and in particular at accumulate economic capital. According to these theories, the choices are guided by the evaluation of economic benefits and costs for individuals, and therefore the differences in wages in different areas are

the key factor triggering movements [55]. However, later empirical studies show that the poorest people from the least-developed countries are unable to undertake the journey towards the richest countries, instead, people of intermediate socio-economical status that have the means to bare the costs are more likely to move in other regions [48].

Today, the neo-classic theories of migration are considered oversimplified and incapable of a complete explanation of migration patterns. One of the first attempt to overcome the difficulties of the neo-classic theories of migration came from Stark [56] who argued that the "agent" of migration, the one who decide to migrate, is not necessarily a single individual, but it can be an entity of a larger scale, such as a household or a group of families bound by an "inter-temporal contractual arrangement" aimed at regulate the costs and the subsequent remittances. In this "new-economics" view, migration is not just a mean to maximize individual utility, but migrant and non-migrant families behave collectively according to a "calculated strategy" in order to diversify their income and consequently minimize their economic risk.

Subsequently, other theories of migration asserted that factors other than the economic ones must be considered to explain thoroughly the behaviour of migrants. In general, a widely adopted view assumes that the process of migration may be epitomize by four factors [57]:

1. **Push factors**: features of the location of origin that "repel" people from it (e.g. scarcity of jobs, political repression, conflicts, environmental degradation);

2. **Pull factors**: features of the location of destination that "attract" people to it (e.g. higher wages, job opportunities, healthier environment, higher perceived quality of life);

3. **Intervening opportunities/obstacles**: features of the locations between origin and destination that may attract the migrants, or may impede the journey (e.g. distance, physical or political obstacles...);

4. **Personal factors**: personal rational and irrational components of the individual decision-making (e.g. age, sex, sensitivity, knowledge and perception of situation at origin and destination...).

On these factors are based some of the best known mathematical models for migration, that will be presented in the next section.

The social theories based on "push-pull" factors consider migration as an essentially individual-household decision-based process, whereas in many cases large-scale mobility and migration patterns are controlled – or at least prompted – by

political or economic entities pursuing their development needs. Mass recruitment of workers for factories, agribusiness and major infrastructure projects marked the colonialism domination of the richest countries on the underdeveloped peripheries. Castles and Miller [48] argue that "migration was as important as military hegemony and control of world trade and investment in keeping the Third World dependent on the First". As such, the political power of countries and the interests of economic entities put constraints on migrants' choices, and consequently on the large-scale patterns. This is the fundamental argument of the "historical-structuralist theory" of migration, that gives prominence to the role of international relations, political economy and the historical causes of mass human movements. These factors are often tricky to express into quantitative terms and also difficult to predict.

An alternative explanation of the migration processes and their causes come form the "migration system theory", that draws attention on the existence of prior relations between countries of origin and destination, such as political influence, trade, shared languages and/or religion. A migration system is constituted by the set of countries from which migrants leave and towards which they go, along with the various types of relations connecting them. This theory integrates the *micro-* and *macro-structures* that trigger human movements [48] in a unifying framework. The macro-structures refer to those factors that regulate the system at large-scale, such as political and economical ties between countries and regions, while the micro-structures are akin to the above mentioned "personal factors". Importantly, these include the social networks of the migrants at both locations of origin and destination that could provide help to further migrants (e.g. by sharing information, organizing travels, finding work). The establishment of such social relations act as a positive feedback force on migration flows that often become self-sustained [58]. In general, the focus on interdependence and reciprocity of different structures emphasize that flows of people can be hardly explained without taking into account other types of relations and flows (e.g. trade flows) that contribute to shape each other.

From the end of the last century, scholars have started to study the role played by the environmental conditions on large-scale human mobility [59, 60, 61]. Both global and local environmental changes may be the cause of increased migration flows, and even the cause of definition of new routes of migration [62, 63]. The influence of climate change, in particular, has ignited discussions because of its inherent uncertainties, the importance of political and social effort and the increasingly frequent environmental adverse events [64, 65, 66, 67]. As all living species on Earth, also humans live and construct their ecological niche co-evolving with the environment [68], and climate change is likely to alter the spatial niche where humans have settled during history (see fig. 1.1). Despite this, we should note that

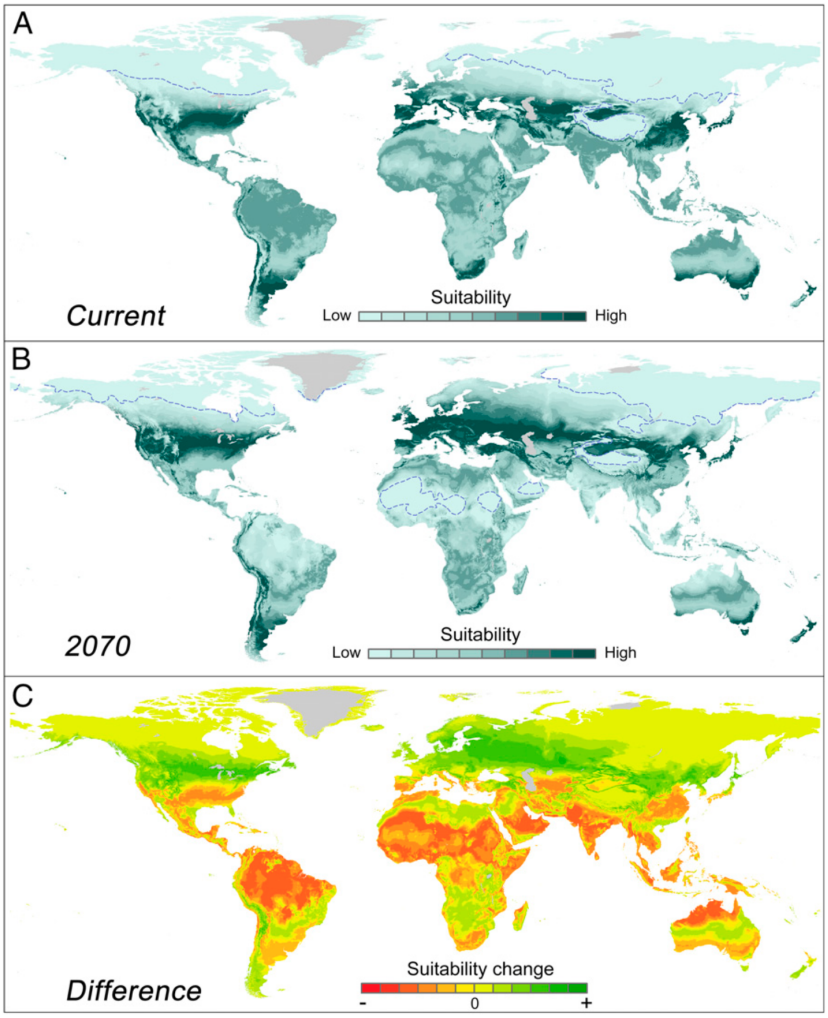


Figure 1.1 – Spatial shift of the human temperature niche – Spatial distribution of the human temperature niche at the current time (**A**) and under a RCP8.5 climate scenario projected in 2070 (**B**). These maps represent the relative human distributions (summed to unity), assuming that humans would be distributed over temperatures following the current population distribution as a function of the mean annual temperature. The dashed lines indicate the 5% percentile of the probability distribution. The panel (**C**) depicts the difference between the maps **B** and **A**, highlighting potential sources (orange) and sinks (green) areas assuming that humans would move in order to maintain the historically stable distribution with respect to mean annual temperature. Figure and caption from [72]

the response of the migration system to the environmental changes is highly non-linear [69, 70], and moreover, forms of mitigation and adaptation strategies, other than migration, are preferably adopted to cope with environmental hazards [71]. Nevertheless, environmental changes are recognised to be direct and/or indirect cause of migration, being one of an assemblage of diverse drivers. The most widely

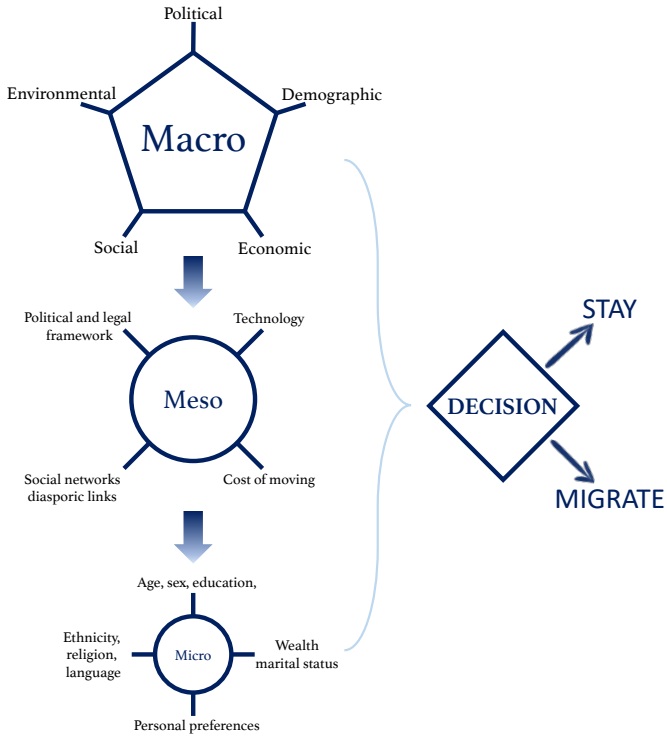


Figure 1.2 – A conceptual framework for the "drivers of migration". Figure adapted from [10].

used conceptual framework in the study of the influence of environmental change on migration and on its concurrent drivers has been proposed by Black in [10]. This framework is pictured in fig. 1.2. Five fundamental macro-structural drivers of migration (the pentagon at the top of the figure) and their mutual relations contribute to shape the conditions of meso-structural, and in turn micro-structural drivers, going from global to individual scale. The macro-structural drivers are constituted by the *economic drivers*, including employment opportunities and income differentials between places; the *political drivers* including the level of security, presence of conflicts, discrimination and persecution, and also other types of policy that have an impact on human movements, such as international agreements, enforced relocation and incentives to immigrants; the *demographic drivers* such as the size and structure of populations in source areas, the presence of diseases, the growth and mortality rates; the *social drivers* including the social relations with other migrants and non migrants at origin and destination, familial or cultural expectations, educational opportunities and cultural practices; the *environmental drivers* of migration are exposure to hazard and availability of ecosystem services. It is important to notice that these drivers do not act as separate compartments, but the interactions between them affect the conditions at lower scales and eventually the

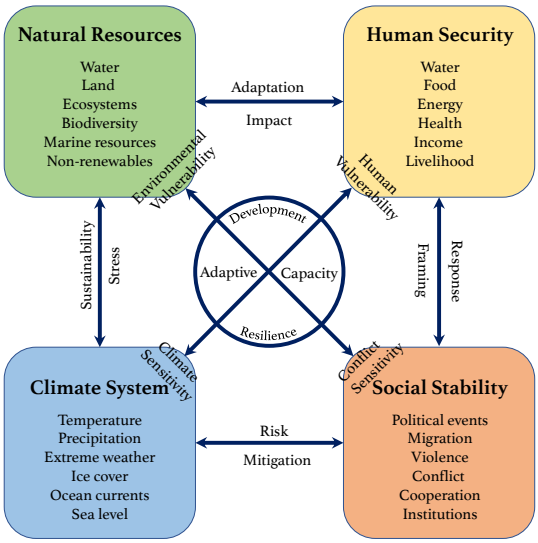


Figure 1.3 – **Analytical framework** of linkages between the climate system, natural resources, human security, and societal stability. Figure adapted from [73].

details of movements. In fact, the actual or perceived differences across space in the drivers, at different scales, eventually affect the choices of migrants. In turn, the multiple interactions and feedbacks between the climate system, natural resources, human security, and societal stability are represented in fig. 1.3 adapted from [73]. These macro and meso scale factors contribute in molding the living conditions of people and consequently their willingness to stay or move in other regions. For instance, various scenarios suggest that in the next century climate change is expected to be an increasingly relevant direct and indirect cause of migration (see e.g. [69, 74]), increasing *climate sensitivity*. Also, scarcity or mismanagement of natural resources is proved to be linked to violence and armed conflicts (*conflict sensitivity*), that in turn prompt human displacement and migration [75].

The climate system itself is characterized by a complex causal structure in which the dynamics of the environmental, social, economical and political systems interact and respond to feedbacks according to a network of interactions. Figure 1.4) from [76]) highlight how this system can be understood as a complex network, and as such, network tools can be used to find central nodes and communities, feedback loops, etc. In chapter 3 we propose a new framework to conduct this kind of analysis on complex networks, that is robust against the uncertainty on the causal structure.

It is known that during history harmful environmental events, and in particular climate changes, induced societal crisis that caused mass human movements [77, 78]. On the one hand, non-catastrophic environmental events, such as

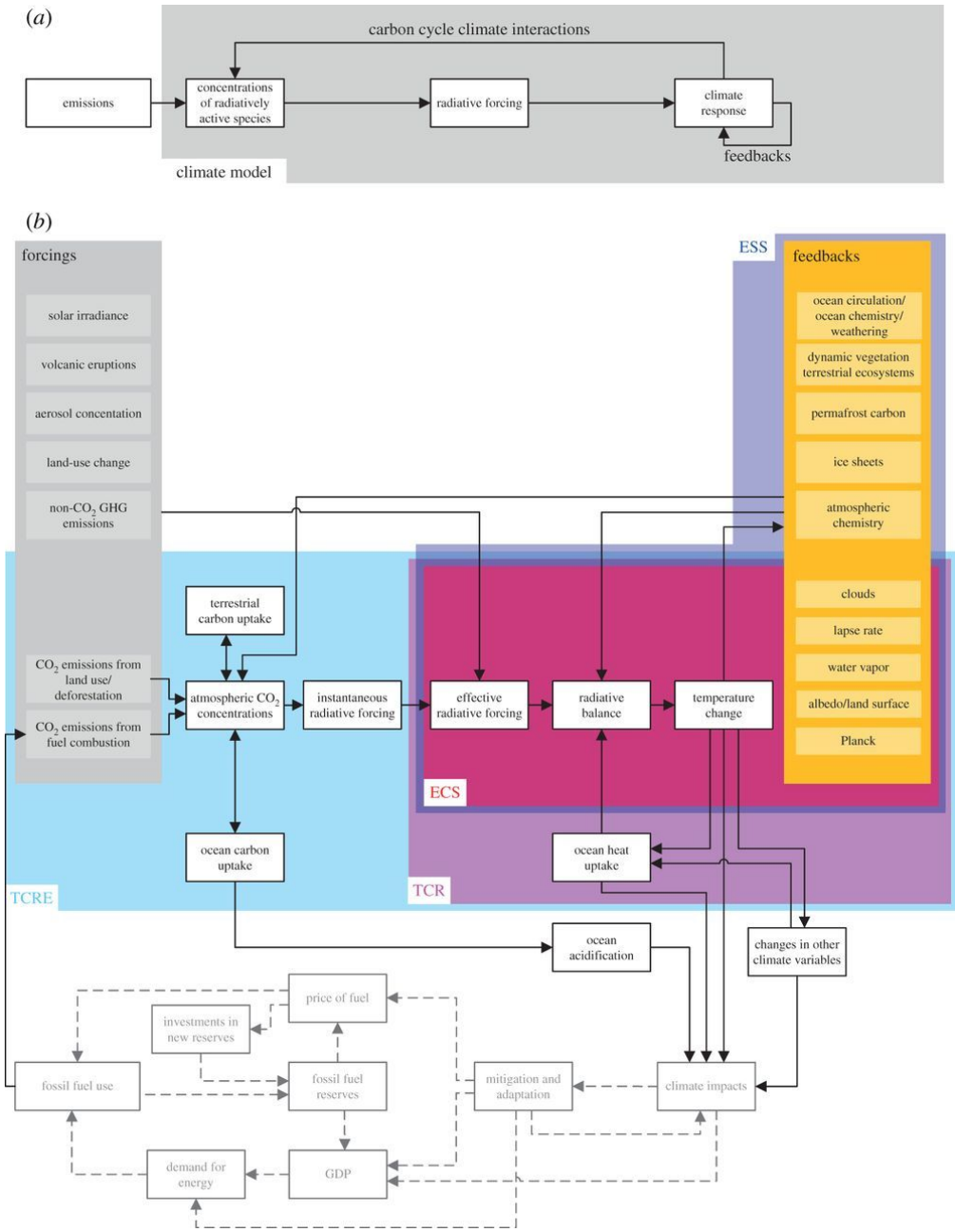


Figure 1.4 – The causal network structure of the climate system – Climate change, through temperature and other variables change will impact socio-economic systems, which finally will feed back on emissions. See text for further discussion. The feedback loops sketched act on different timescales. Reprinted with permissions from [76]

sea level rise or increasing temperatures, are potential slow-onset drivers of migration [79], albeit, as previously said, a single cause is hardly sufficient to explain thoroughly the decision to move [48, 65]. On the other hand, extreme environmental events may impact human mobility by triggering migration, displacement, or forced immobility [80], especially in those contexts where people are already exposed to risks to life, health and well-being. "Environmental migration"[1] is a full-fledged adaptation strategy against adverse events, such as severe droughts, floods or earthquakes [85, 86, 87, 88], whereas "displacement" is a particular form of migration, in which individuals are forced to move against their will. Movements can be permanent or temporary, depending on the severity of the event and on other external conditions; for instance, people may seek shelter before, during or after a flood or a hurricane. The displacement of people caused by such extreme events is the topic of the chapter 5, in which we present a quantitative and qualitative analysis for a couple of case studies. We estimate a gravity model combined with penalized regression to identify the drivers of mobility patterns.

In the following section, we briefly describe the mathematical models – including the gravity model – that have been used so far to understand the forcing drivers of migration and to predict migration patterns at different scales.

[1]Even though environmental migration is a fairly clear fact, the definition of "environmental migrant" is a debated topic both by human rights scholars and political, social and environmental scientists [65, 81, 82, 83, 84]

Chapter 2

Mathematics for complexity and causality in human mobility

> Nothing is more important than to see the sources of invention which are, in my opinion more interesting than the inventions themselves.
>
> G. W. Leibniz

2.1 Mathematical models for human mobility

The mathematical description of human mobility can be divided into two main categories: individual level models, and population level models. The first class aims at describe the movements of single individuals in space and time. To this class belong the random walk models, that are inherently stochastic, and the agent-based models that can be both stochastic or deterministic. The typical result of this kind of models is the spatial distribution of individuals or a probability distribution that provides the probability of an individuals to be in a certain position. The population level models, instead, describe the aggregate mobility of many individuals by estimating the number of travelers, or the probability of a trip between any two spatial locations. The result is usually an Origin–Destination (OD) matrix, whose entries correspond to the flow of people between two locations or the probability of such a flow.

Random Walk The simplest mathematical prototype used to model individual human mobility is the random walk. A random walk is mathematically defined as a path formed by successive random steps in a n-dimensional space [89]. In one

dimension, the process $\{S_n, n \geq 0\}$ is called a random walk process if

$$S_0 = 0, S_n = \sum_1^n X_i, n \geq 1 \tag{2.1}$$

and $X_1, X_2, \ldots$ are independent and identically distributed random variables representing the "steps", with $E\left[|X_t|\right] < \infty$. A particular type of random walk is the Brownian motion, originally defined by the botanist Robert Brown to describe the motion of particles in a liquid or gas [90]. In this case, we have a step increment equal to Δx during a time increment Δt so that the position $X(t)$ of the walker at time t is defined as

$$X(t) = \Delta x \left(X_1 + \cdots + X_{[t/\Delta t]}\right) \tag{2.2}$$

where

$$X_1 = \begin{cases} +1 & \text{if the } i \text{ th step of length } \Delta x \text{ is to the right} \\ -1 & \text{if it is to the left,} \end{cases}$$

and where the X are assumed independent with

$$P\left\{X_1 = 1\right\} = P\left\{X_t = -1\right\} = \frac{1}{2}.$$

Since $E\left[X_t\right] = 0, \mathrm{Var}\left(X_t\right) = E\left[X_t^2\right] = 1$, we see from eq. (2.2) that

$$E[X(t)] = 0$$
$$\mathrm{Var}(X(t)) = (\Delta x)^2 \left[\frac{t}{\Delta t}\right]$$

The variance scale linearly with time, while the mean remains constant and the probability to observe a displacement of magnitude $X(t)$ from the origin location after a time t is Gaussian distributed with mean zero and variance proportional to t. This is a consequence of the Central Limit Theorem for $t \to \infty$ in eq. (2.2). It can be proven that this process is continuous, and is a limit case of the random walk defined in eq. (2.1) [90].

The simple random walk and the Brownian motion permit to identify a typical length scale for the displacements, since the Mean Square Displacement (MSD) – defined as the second moment of $X(t)$ – scale linearly with time at any spatial dimension and the mean is finite. This kind of process is called "diffusive". In many real world application, no typical length scale can be identified and these processes would fail in describe the nature of motion. When MSD (or analogous quantities) deviates from linearity, the process become *sub-diffusive* or *super-diffusive*. An example of super-diffusive process used to describe movements in both humans

and animals is the Lévy Flights [91]. This process is composed of a series of small displacements, interspersed occasionally by a very large displacement. The probability density function of the step increment is

$$p(\Delta x) = \frac{1}{\Delta x^{1+\beta}} \tag{2.3}$$

with $0 < \beta < 2$. The occasional large displacement of the Lévy flights happen in the same time of the normal ones, that may be unrealistic to describe human movements.

In some cases, the trajectories of human movements proved to be far from being random walks [92], showing a high degree of temporal and spatial regularity, where individuals follow a time-independent characteristic travel distance and a significant probability to return to a few highly frequented locations [2]. In particular, mobility data show that the frequency of return at different locations follows a Zipf law of the type

$$f_k \sim k^{-\alpha} \tag{2.4}$$

where k is the rank of the location according to the frequency of the visit and α the characteristic scale parameter. A model that takes into account both the "exploration" behaviour of a random walk, and the "preferential returns" observed in humans was proposed in [92]. The model describe the movements by using two complementary probabilities:

$$P_{new} = \rho S^{-\gamma} \tag{2.5}$$

$$P_{ret} = 1 - P_{new} = 1 - \rho S^{-\gamma} \tag{2.6}$$

where P_{new} defines the probability for an individual to move to a previously unvisited location, while P_{ret} is the probability to return to a previously visited location. The two terms ρ and γ are the parameters of the model both of which control the tendency of the walkers to either explore a new location or returning to a previously visited location in the next step. The numerical values of these two parameters are determined using empirical data. The value of S increases by one for each newly visited location.

A stochastic model intimately related to the random walks is the Markov chain. This model has been used to describe both large-scale mobility [93, 94, 95] and local displacement of refugees [96]. The basic idea of this model is to assign a probability of transition from origin to destination that is a function of utility and costs of moving. People move from origin to destination until the gain in utility is exactly offset by the cost of migration. If the Markov chain is appropriately defined (see e.g. [97] for a detailed exposition of the model), the resulting stationary distribution stands for the population distribution over the locations.

Agent Based Models An Agent-Based Model (ABM) is a computational model that simulates the behaviour of autonomous agents within an environment. The agents are autonomous, individual elements with prescribed properties capable of making decisions, take actions and responds to stimuli, by interacting with other agents and with the environment. The environment itself can have dynamical properties and its components can interact with each other. The micro-scale behaviour of the agents is typically governed by simple decision-making and logical rules. From the perspective of complex systems, ABMs are used to study the emergent collective properties of complex adaptive systems by modelling the functions of the individual components (see e.g. [98, 99, 100] for an overview of the topic).

In the specific case of human mobility, as said in the previous section, decisions are multi-causal and conditioned on individual preferences. Therefore, agent-based modelling offers a capable method to simulate the autonomous decision making process, accounting for unobserved differences between individuals. People are attracted by locations with higher utility or value. In modern ABMs the outcome of migration is uncertain, since utility and value are random variables and also the internal state of agents (such as the marital status, job opportunities, or retirement) may comprise a stochastic component. Movements are triggered by the expected utility or the expected value. Besides the commuters/migrants, other agents can be modelled, such as other individuals, friends and institutions that may influence the decisions of the moving agents on the destination to choose or the timing and means of movements.

The aim of an agent-based model is to uncover causal mechanisms that lead to human movements [101]. ABMs stands as a counterpart to "equation based models" since they lack equations which govern the overall social structure at the macro-level [102]. The micro-level decision making not necessarily reflects the (quasi-)rational behaviour of people, but may be defined in terms of simpler rules, such as laws of attractions and agents walking at random [103], or empirical rule-determined decisions as in [104]. Other examples are the models based on expected utility maximisation [105], on psycho-social and cognitive models (see e.g. the theory of planned behaviour applied to environmentally induced migration in [106]) and models halfway between the economic ABMs the empirical models such as in [107].

Agent-Based Models can be used to unravel the micro-structural drivers of migration, and also to predict future movements in order to support decision making (see e.g. [108] for a case study on refugees). Nonetheless, there is an ongoing debate on whether prediction should be a major purpose for ABMs [109]. Specifically for migration studies, the definition of decision-making theories and the selection of empirical evidence for model validation are major challenges for this type of model [101].

Random Utility Models Another class of models commonly used to characterize decision-making in economics, transportation engineering and human mobility research, is that of Random Utility Models (RUM). These models are used to describe the behavioural process that leads agents to choose among a set of alternatives [110] – in our case, a set of destinations. The choice set is assumed to be finite, exhaustive, and the elements must be mutually exclusive. The decision-making process is described assuming a causal perspective in that the behavioural process is defined as a function of external factors that determine the choice of an agent. The external factors are divided into observable deterministic factors x and non-observable stochastic disturbances ε. The choice of the agent can be defined as $y = h(x, \varepsilon)$, where y is the alternative h is the behavioural process. Given the uncertainty on ε, the choice of the agents must be expressed in probabilistic terms as

$$\begin{aligned} P(y \mid x) &= \text{Prob}(\varepsilon \mid h(x, \varepsilon) = y) \\ &= \text{Prob}(I[h(x, \varepsilon) = y] = 1) \\ &= \int I[h(x, \varepsilon) = y] f(\varepsilon) d\varepsilon \end{aligned} \tag{2.7}$$

where $I[\cdot]$ is the indicator function, and $P(y \mid x)$ is the probability that, with x given, the stochastic component is such that the behavioural process leads to the choice of the alternative y. Specifically, this probability can be interpreted as the expected value of the indicator function over all the possible value of the stochastic disturbance.

Equation (2.7) is valid in general, for any specified behavioural process, and is the base of any the *discrete choice model.* The behavioural process of the RUMs, are defined according to an utility maximization criterion. The utility V_{ni} that the decision-maker n assigns to the alternative i is a function of the attributes x_{ni} of the alternative as perceived by n, and of the attributes s_n of the decision-maker it self, more precisely

$$V_{ni} = V(x_{ni}, s_n) \tag{2.8}$$

where V_{ni} is called "representative utility", and corresponds to the deterministic factor observable by the researcher. The function $V(\cdot)$ can have any form, and in general depends on parameters that are estimated with data. The utility that the researcher actually assign to the alternative is biased by its partial knowledge of the representative utility, and so we should define the actual utility as

$$U_{nj} = V_{nj} + \varepsilon_{nj}. \tag{2.9}$$

The probability that the agent n choose the alternative i by maximizing his utility

is thus given by

$$\begin{aligned} P_{ni} &= \text{Prob}\left(U_{ni} > U_{nj} \forall j \neq i\right) \\ &= \text{Prob}\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall j \neq i\right) \\ &= \text{Prob}\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i\right) \\ &= \int_{\varepsilon} I\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i\right) f\left(\varepsilon_n\right) d\varepsilon_n \end{aligned} \tag{2.10}$$

where $f(\varepsilon_n)$ is the joint density of the random vector $\varepsilon_n' = \langle \varepsilon_{n1}, \ldots, \varepsilon_{nJ} \rangle$.

Many different flavours of RUMs can be implemented depending on the form of the distribution $f(\cdot)$. For example, using the Gumbel distribution[1] one obtains the so called *logit* model [111], that has the desirable characteristic of being analytically tractable, since the integral for the choice probability (eq. (2.10)) has a closed form solution, and so is one of the most used form of RUM. The logit, however exhibits *independence from irrelevant alternatives* that is, adding or removing an alternative has no effect on the ranking of the original set of alternatives. In other words, the logit model cannot take into account any correlation between the stochastic disturbances of the alternatives that may be present in the data. If this is the case, other types of model can be chosen, for example, the *nested logit* that assume $f(\cdot)$ to be a Generalized Extreme Value distribution (see [110] for a complete survey on the different types of RUMs). This distribution allows for correlations over alternatives, and it is used for example in [112] to cope with multilateral resistance to migration, as explained in the next paragraph dedicated to the gravity model.

Gravity Model The gravity model is the workhorse of human mobility literature since its formulation. Ravenstein [53, 54] laid the theoretical foundations of the model, as already discussed in section 1.2. More than sixty years later, G.K. Zipf [113] gave the first[2] formal expression of the model by showing that the number of persons that move between any pair of locations in the United States

[1]The Gumbel distribution corresponds to the Generalized Extreme Value distribution Type-I, and is used to model the distribution of the maximum (or the minimum) of a series of samples from other distributions. By assuming that $f(\cdot)$ is a Gumbel of the form $f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}}$ and that the ε are independent, eq. (2.10) becomes

$$P_{ni} = \int \left(\prod_{j \neq i} e^{-e^{-(\varepsilon_{ni} + V_{ni} - V_{nj})}} \right) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

that is the logit choice probability.

[2]Earlier attempts to explain the patterns of human spatial interaction flows with a mathematical model can be found in [114, 115, 116])

are proportionate to the ratio, $P_1 \cdot P_2/D$, where P_1 and P_2 are the population at the two locations that are separated by the shortest transportation distance, D. The model has been successfully adopted also by economists as an explanation of bilateral trade flows, following the mathematical expression by Tinbergen in 1962 [117].

The gravity model, needless to say, is developed by analogy with Newton's law of universal gravitation, that takes the form:

$$\mathbf{F}_{21} = -G\frac{m_1 m_2}{|\mathbf{r}_{21}|^2}\hat{\mathbf{r}}_{21} \tag{2.11}$$

where $\mathbf{F}_{21}$ is the force applied on object 2 exerted by object 1, G is the gravitational constant, m_1 and m_2 are respectively the masses of objects 1 and 2, $|\mathbf{r}_{21}| = |\mathbf{r}_2 - \mathbf{r}_1|$ is the distance between objects 1 and 2 , and $\hat{\mathbf{r}}_{21} \stackrel{\text{def}}{=} \frac{\mathbf{r}_2-\mathbf{r}_1}{|\mathbf{r}_2-\mathbf{r}_1|}$ is the unit vector from object 1 to object 2.

In migration and economic studies, the formula in eq. (2.11) is simplified and the terms acquire a different meaning: in the place of masses, various indices of "attractiveness" are considered, such as the population, the income or the jobs opportunities in two locations; the squared distance $|\mathbf{r}_{21}|^2$ is replaced by a more general function that represents the cost of moving. In the original model by Zipf the attractiveness of the locations is assumed to be proportional to the population and the cost is proportional to the traveled distance. We can rewrite the model as follows:

$$T_{ij} \propto \frac{P_i P_j}{r_{ij}} \tag{2.12}$$

where P_i and P_j are the respective populations and r_{ij} the distance between i and j. The term on the left hand side is the flow of people between locations i and j, and represents the term (i, j) of the so called "Origin-Destination matrix" (OD). This equation is a simplistic expression of the gravity model, that has at least one evident deficiency: if the populations of both the locations i and j are doubled, the flow T_{ij} would quadruple, whereas it is expected to just double. More precisely, the model should satisfy the following constraints:

$$\sum_j T_{ij} = P_i \tag{2.13}$$

$$\sum_i T_{ij} = P_j \tag{2.14}$$

that means the total outflow from a location must be equal to the sum of the observed outflows, or in other words, that the row and column sums of the OD matrix should be equal to the total outflow and inflow, respectively. Equation (2.12) does not satisfy these constraints, and the usual remedy is to multiply each variable

with a "balancing" constant [118, 119]. Moreover, the cost factor at the denominator of eq. (2.12) can be considered as a generic "friction function" of the distance between locations. Therefore, a more general formulation reads:

$$T_{ij} = A_i B_j P_i P_j f\left(r_{ij}\right) \tag{2.15}$$

The form of the friction function $f\left(r_{ij}\right)$ can be identified by means of entropy maximization [118]. In order for the method to work, another constraint should be added on the total cost C of the travels, namely:

$$\sum_i \sum_j T_{ij} c_{ij} = C \tag{2.16}$$

where c_{ij} is a generalized cost. The basic assumption of the method is that the probability distribution of the T_{ij} is proportional to the number of states of the system that satisfies the constraints in eqs. (2.13) and (2.16). Maximize the entropy of the system means to find the configuration of the OD matrix that maximizes the number of possible configurations of trips associated with it. Let us first consider the first element of the OD matrix, indexed with $(1,1)$. The number of ways in which $T_{1,1}$ individuals (taking a self-loop-trip) can be selected out of the total $T = \sum_{ij} T_{ij}$ individuals is given by $\binom{T}{T_{1,1}} = \frac{T!}{T_{1,1}(T-T_{1,1})}$. Similarly, for the OD element $(1,2)$ the possibilities are $\binom{T}{T_{1,2}}$. The number of configurations of the whole OD matrix is then given by

$$\Omega(T_{ij}) = \frac{T!}{T_{11}!\,(T-T_{11})!} \frac{(T-T_{11})!}{T_{12}!\,(T-T_{11}-T_{12})!} \cdots = \frac{T!}{\prod_{ij} T_{ij}!} \tag{2.17}$$

The total number of possible states is then equal to $\sum_{(i,j)\in\mathcal{C}} \Omega\left(T_{ij}\right)$, where the $\mathcal{C}$ is the set of trips that satisfies the constraints. The configuration that maximizes the number of distinct arrangements of individuals giving rise to a particular OD matrix is obtained by maximizing $W\left(T_{ij}\right)$ under the defined constraints. The maximum can be found using the method of the Lagrange multipliers. The Lagrangian expression is defined by

$$M = \log w + \sum_i \lambda_i^{(1)} \left(P_i - \sum_j T_{ij}\right) + \sum_j \lambda_j^{(2)} \left(P_j - \sum_i T_{ij}\right) + \\ + \beta \left(C - \sum_i \sum_j T_{ij} c_{ij}\right) \tag{2.18}$$

where $\lambda_i^{(1)}, \lambda_j^{(2)}$ and β are the Lagrangian multipliers. The most probable distribution of trips is given by the T_{ij} that maximizes M being the solution to

$$\frac{\partial M}{\partial T_{ij}} = 0. \tag{2.19}$$

It can be shown, after calculation, that the solution is given by

$$T_{ij} = A_i B_j O_i D_j \exp\left(-\beta c_{ij}\right). \tag{2.20}$$

The solution to the maximum entropy problem returns the gravity equation in the form of eq. (2.15), where the functional form of the friction is $\exp\left(-\beta c_{ij}\right)$. The parameter β can be calibrated using empirical data. Different formulations of the constraints lead to different friction functions, as extensively elaborated in [119] and [120].

Frequently, the total outflow $O_i = \sum_j T_i j$ originating at location i may be hold constant and equal to the observed total outflow, while the gravity model is then used to estimate each D_j. This variant of the gravity model is called *singly-constrained* gravity model.

Even though the structural analogy to the universal gravity law is useful to intuition, the gravity model is no way *universal*, in that there is no set of parameters for which the eq. (2.20) is valid for any arbitrary set of observations; nevertheless, in practical implementations of the model the parameters can be calibrated to fit data in different settings. Several controlling/dummy variables may be included in the attractiveness terms along with the population, in order to take into account the economic conditions (e.g. GDP, income), the share of a common language and/or religion, the colonial past, the environmental conditions, the number of job opportunities, among others (see e.g. [121, 122]). Also the form of the friction function can be adapted to achieve the best fit [123, 124, 125]. The presence of numerous regressors and parameters require large dataset for calibration and increases the risk of overfitting and multicollinearity (see example in section 1.2). The model calibration can be carried out using different methods depending on the model structure. The most simple way is by using Ordinary Least Square (OLS) after log-linearize eq. (2.15). However, this approach may lead to an inconsistent and biased estimation of the parameters due to the concurrence of heteroskedasticity (highly probable in applications to mobility data) and non-linearity. To overcome this issue, Silva and Tenreyro [126] suggested to use the Poisson-Maximum Likelihood estimation (see also [127, 128] for details on the technique) that is robust to different patterns of heteroskedasticity and also provides a way to deal with zero values of the dependent variable (zero flow route) that would conversely makes log-linearization unfeasible. Note that other methods for model calibration are available, such as Non-Linear Least Squares and Structurally Iterated Least Square (see [129, 130, 131] for an overview).

The complex structure of the migration systems gives rise to a larking driver of migration that is called called "multilateral resistance". The recognition of such a driver comes from the economic literature[3] to explain the influence exerted by

[3]First identified by Anderson in [132] and then expanded by himself in [133]. The effect is

all the possible alternative destinations on the bilateral flow of goods. To capture the multilateral trade resistance Feenstra [134] and Redding and Venables [135] proposed to include importer and exporter fixed effects in the regression model. In 2013, Bertoli and Moraga [112] gave a definition of multilateral resistance to migration, using a RUM model with a nested stochastic component of utility that follows origin-specific patterns of correlation across alternative destinations (see also [129] for a survey on the topic). Recently, Simini et al. [136] proposed an interpretation of the "flow generation problem" as a classification problem, guided by a singly-constrained gravity model. They use geographical features to train a deep neural network, namely the Deep Gravity model, to estimate mobility flows between locations identified by a set of geographical polygons called tassellation. The Deep Gravity model improved on the simple gravity model especially in densely populated regions.

Intervening Opportunities Model For many years, the most valuable alternative to the gravity model was the intervening opportunities model introduced by Stouffer in 1940 [137]. If the gravity model contains the explicit influence of the distance on the mobility flows, the intervening opportunities model assumes that the relationship between mobility and distance is not necessary. Instead, it states that "the number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities". More formally, let Δy be the number of persons moving from the origin to a surrounding annulus centered on the origin of width Δs. The annulus start at distance $s - \frac{1}{2}\Delta s$ and ends at $s + \frac{1}{2}\Delta s$ (see fig. 2.1). The intervening opportunities, are the cumulated number of opportunities between origin and distance s and we identify them with x. Let also Δx being the opportunities contained in the annulus centered at distance s and wide Δs. Then, the original formulation of the model is expressed as follows:

$$\frac{\Delta y}{\Delta s} = \frac{a}{x}\frac{\Delta x}{\Delta s} \tag{2.21}$$

where a is a proportionality constant to be estimated. This model postulates a relation between mobility and opportunities, rather than between mobility and distance. The precise definition of opportunities is a problem specific problem, that depend on the social phenomena under investigation. Interestingly, defining $x = f(s)$ as the cumulated opportunities within a circle of radius s, with f assumed continuous and differentiable, eq. (2.21) can be rewritten as

$$\frac{\mathrm{d}y}{\mathrm{d}s} = \frac{a}{x}\frac{\mathrm{d}x}{\mathrm{d}s} \tag{2.22}$$

called "multilateral trade resistance" in the economic context.

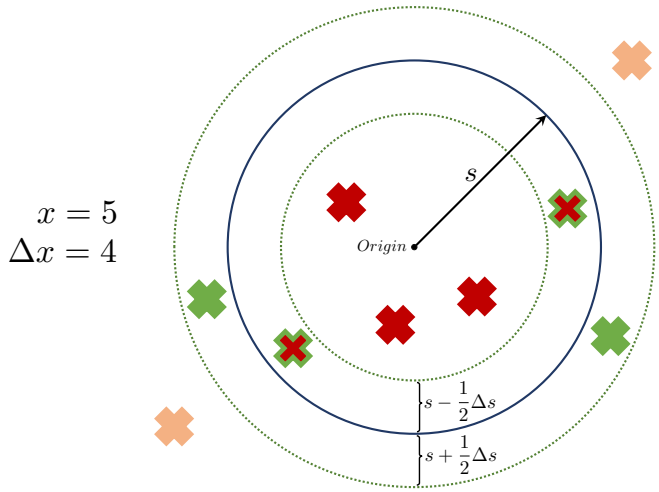


Figure 2.1 – Intervening opportunities representation. The individuals move from the origin to the circular band centred at distance s. The opportunities are represented as colored crosses. The intervening opportunities x are represented as *red* crosses, contained within the radius s. The crosses inside Δs are *green*. The *red-green* crosses are both inside s and Δs.

that after integration gives

$$y = a \log f(s) + c \tag{2.23}$$

that provides the number of movers that starting from the origin choose as destination any location within the circle of radius s. Stouffer justifies the form of eqs. (2.21) and (2.23) as being consistent with the decreasing knowledge of the opportunities with distance, that is in turn compatible with the Weber-Fechner law of psychophysics [138], which relates stimuli and perceptions logarithmically.

Another formulation given in literature, attributed to Schneider [139, 140], rely on the assumption that every location of destination has a stated constant probability of being accepted if it is considered. Than, it follows that

$$\mathrm{d}P = L\,[1 - P(V)]\,\mathrm{d}V \tag{2.24}$$

where $\mathrm{d}P$ is the probability that an individual goes from origin to $\mathrm{d}V$ possible destinations; $P(V)$ is the probability that a trip will terminate by the time V possible destinations are considered; V defines the possible destinations already considered; L is the constant probability of a possible destination being accepted if it is considered.

The solution to the eq. (2.24) with $P(0) = 0$ is

$$P(V) = 1 - \exp(-LV) \tag{2.25}$$

The (unnormalized) expected flow T_{ij} from location i to location j is equal to the number of individuals O_i starting from the origin, multiplied by the probability

of a trip from i to j:

$$T_{ij} = O_i \left[P\left(V_{j+1}\right) - P\left(V_j\right) \right] \tag{2.26}$$
$$= O_i \left[\exp\left(-LV_j\right) - \exp\left(-LV_{j+1}\right) \right] \tag{2.27}$$

The value of the parameter L (which in general varies for different trips) can be estimated from data. Comparisons between the gravity and the intervening opportunities models show that both model are comparable in terms of performances in practical applications [141, 142, 143]. Although other variants of the model has been proposed (see e.g. [144, 145, 146, 147]), the intervening opportunities model did not have the same popularity of other models, probably because of the lack of research effort into the implementation and calibration of the model and the ease of understanding of the gravity model and simplicity of use of the radiation model

Radiation Model We introduce here the radiation model, that will be illustrated in more details in chapter 6. As mentioned before, during the XX century the gravity model has dominated human mobility models by virtue of its simplicity and good performances. Still, it has some recognized disadvantages: the gravity model cannot explain the direction of the flows between a pair of locations, since the attraction is reciprocal; the original model is completely deterministic and does not allow uncertainty to be quantified and also, the number of parameters to be estimated in order to obtain a good agreement with data can be very large with respect to the data available for calibration. In 2011, Simini et al [19] proposed a novel mathematiacal model similar in spirit to the intervening opportunities model, but inspired by a the physical theory of radiation. The authors assume that every location is a source and receiver of *identical and independent* particles, and define a parameter-free emission/absorption process that allows the probability of a trip to be estimated just from the spatial distribution of population. Specifically, each location is randomly assigned a value of absorption threshold and absorbance, and the particle is absorbed by the closest location whose absorbance is greater than its absorption threshold. The probability of a trip from location i to location j is thus given by:

$$P\left(1 \mid m_i, n_j, s_{ij}\right) = \frac{m_i n_j}{\left(m_i + s_{ij}\right)\left(m_i + n_j + s_{ij}\right)} \tag{2.28}$$

where m_i and n_j are the population at origin and destination respectively, and s_{ij} is the population living within a circle of radius equal to the distance between i and j (see section 2.1). The model depends only on the populations that are considered as a proxy for opportunities and any other driver of mobility. The model implicitly include also the influence of distance, given by the term s_{ij}.

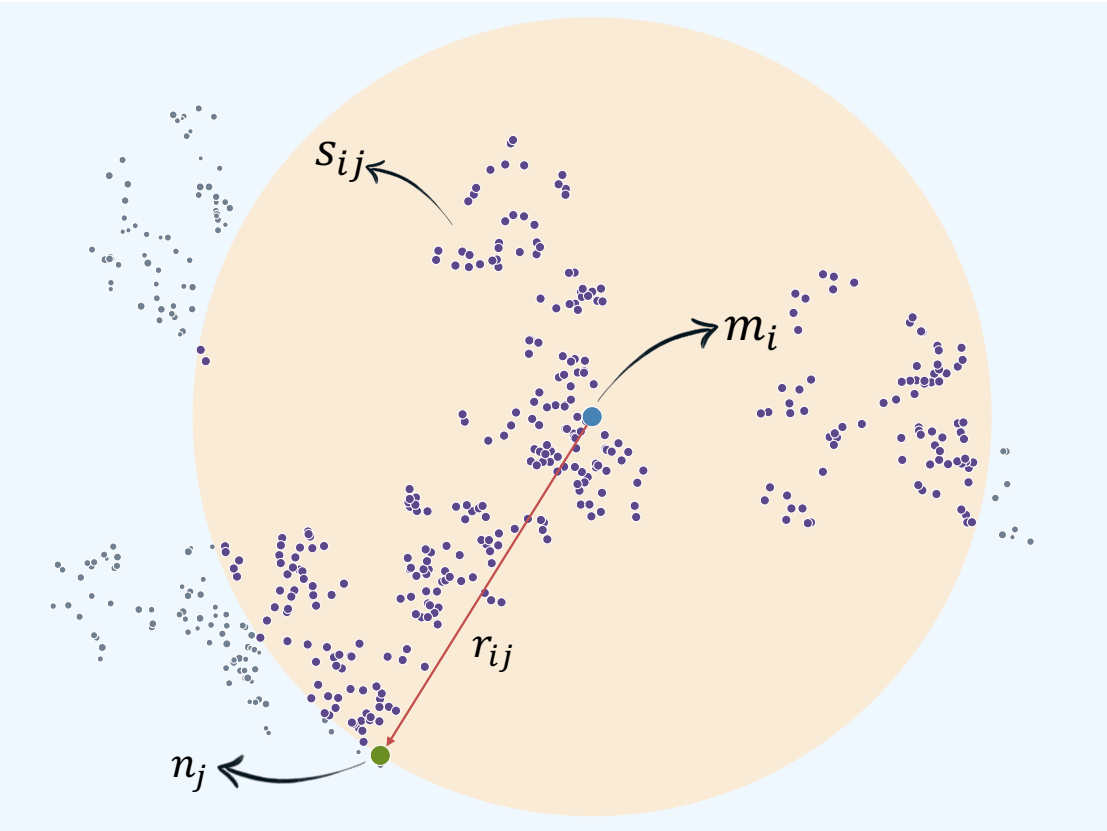


Figure 2.2 – Representation of the terms in the Radiation model – The figure shows a hypothetical spatial distribution of locations. Locations of origin and destination with population m_i and n_j are highlighted in blue and green respectively. The total amount of population inside the circle r_{ij} is denoted with s_{ij}.

The radiation model has been shown to outperform the gravity model in some instances [19]. However, both the models have been shown to poorly describe human mobility in those cases in which the attractiveness is not given just by differences in wages, GDP, populations and related proxies, for example in low income countries [148]. In chapter 6 we propose a generalization of the model in eq. (2.28) that allows other factors to be included as drivers of mobility patterns.

Model of the structure The structural features of a complex system and their interplay with the components dynamics have a direct effect on its collective behaviour, therefore the structural characterization of a system is a fundamental step in understanding the system functioning. Network science has recently inspired scholars to analyse and interpret the very spatial structure of large-scale human mobility. The first proper network analysis of the migration system is attributed

to Tranos (2012) [149], who studied centralities, clusterization and the topological evolution over time using Multiple Regression Quadratic Assignment Procedure (MRQAP). Similarly, in [150] the authors studied the global spatio-temporal patterns of human mobility and proved that the small-world effects has increased over time. Exponential random graph model has been used in [151] to determine geographic, demographic, economic, religious, linguistic as well as historical factors of migration between countries, founding results in line with gravity models. Sgrignoli et al (2015) [152] studied the relation between migration and trade networks, and measured the effect of migration on international trade, while controlling for network interdependences. Belyi et al. (2017) [153] applied multilayer networks to study different kinds of mobility (short-term, long-term, etc.), and they examined the multilayer communities and compared outcomes with those from other existing international connections (e.g. language similarities, present or former colonial relations as well as trade networks), finding relations not detectable from the single-layer networks.

2.2 Mathematical models for causal inference

The causal relations between the components of a complex system and therefore the structure of the system itself, can be quantitatively determined using various methods from statistics, information-theory and causal inference. Each method is based on a specific definition of "causality", from which different characterization of the structure can be deduced. The specific method to infer casual relations between variables, depends on the type of system under study and in particular on the type and quality of the observations. Here we introduce some of those methods used with time series data, the typical observations collected from complex dynamical systems.

In the everyday life, we seem to have a clear intuition of what "causality" means, but a formal and ultimate definition of causality is still argument of debate for scientists and philosophers [154, 155, 156]. The debate is a long-standing one, considering that already Plato argues that causality is not even a part of the physical reality, but rather something that transcend the physical world whereas causal relations are due to the strength of the participation that two events have with a common idea [157]. Similarly, the mathematician Karl Pearson believed that causality cannot be defined in an absolute sense, but rather the universe is "a sum of phenomena, some of which are more, others less closely contingent on each other" [158]. Many other philosophers studied the problem of causality during history, from Aristotle to Hume [159] and Carl Gustav Jung [160]. Recently also many scientists gave their contribution to this problem, considering causality as a physical process that can be studied through mathematics [161, 162, 43]. In this

sense, we introduce some operative definitions of causality that are somehow unrelated to the philosophical interpretations, or rather, the philosophical definition and interpretation of the methods is just that which the methods themself convey.

One of the most remarkable methods of causal inference is due to Judea Pearl [161]. He proposed techniques to infer casual relations from data by introducing a novel mathematical formalism called *do-calculus* to construct rigorous causal models. In general terms, causal inference assesses which causal relations exist and which do not exist by evaluating the effect of perturbations (interventions) that do alter the observations with respect to a counterfactual reality. This framework requires at least well defined causal hypotheses and a causal model of the system under study. Recent works proposed similar approaches to build causal models also using time-series data from dynamical systems (see e.g. [163, 164]). In section 4.4 we describe in detail the *Causal-Impact* model [165] that we used to determine the causal relation between a change in human activity variables (such as human mobility) due to COVID19 lockdown and environmental conditions, using time series data. Another school of thoughts, alternative and auxiliary to the one of J. Pearl on causal inference, is led by A. Gelman [166] and centers around *hierarchical/multilevel models*. This type of models are regression models in which the parameters are random variables with an associated probability distribution (as in most Bayesian settings) that is a stochastic model for the parameters. This "second-level" model has parameters of its own, usually called *hyper-parameters*, which are also estimated from data. From these models the effect of interventions on a population can be estimated being the difference between the predicted outcome conditional on the treatment [167].

In many cases, especially in complex dynamical systems, there is no need to define a causal relation using a "counterfactual" or "intervention" argument, or even, it would not be possible to alter the system dynamics for various reasons. In the following we present a non-exhaustive list of methods for causal analysis, in particular we introduce some of the so called "exploratory causal" tools (as opposed to "confirmatory" [168]) that we used in the following chapters to detect causal relations between dynamical units of complex systems.

Correlations Until recently, causality was virtually prohibited from statistics, on the wake of the old saying "correlation is not causaltion" [156]. Various types of correlation has been used though to get hints about causal structures in brain networks [169], economics [170], systems biology [171], among others.

The first appearance of correlation in mathematics is attributed to A. Bravais for a work of 1844 [172]. The concept was then investigated empirically by F. Galton [173] and named after K. Pearson who further developed the theory [174]. Given two discrete random variables X and Y with probability distri-

bution $P(X = x_i) = p_i$ and $P(Y = y_j) = q_j$ with mean μ_x and μ_y and standard deviation σ_x and σ_y respectively, the Pearson correlation coefficient is defined as

$$\rho_{X,Y} = \frac{\mathrm{E}\left[(X - \mu_X)(Y - \mu_Y)\right]}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{2.29}$$

whereas given the sample mean $\bar{x} = \frac{1}{T}\sum_{i=1}^{T} x_i$ the sample correlation coefficient reads

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2.30}$$

The Pearson correlation coefficient is a measure of the *linear* correlation between two dataset. It can be interpreted as the normalized covariance between the two variables X and Y, that always has a value between -1 and 1. Since this coefficient can quantify just the linear correlation of variables, it ignores other types of relationship or correlation. Some variants of the Person coefficient exist, which can also takes into account mild non-linear relationship. The Spearman's rank correlation coefficient, for instance, gives a measure of the monotonicity of the relation between X and Y, and is equal to the Pearson correlation between the *rank* values of these two variables. Similarly, the Kendall's rank correlation coefficient measures the similarity in the ranking of the two dataset.

For the analysis of complex systems, and in general of dynamical entities which have an evolution over time, the temporal component plays a central role in defining the relation between the units of the system. In fact, an important property – also considered as an axiom of causality [175] – states that *causes precede effects*. Consequently, the correlation coefficient in eq. (2.29) can be adapted to take into account time progression:

$$\rho_{xy}^{l} = \frac{E\left[(X_t - \mu_X)(Y_{t-l} - \mu_Y)\right]}{\sigma_X \sigma_Y} \tag{2.31}$$

This is the definition of the (lagged) cross-correlation coefficient, where l is a time-lag between the two time series $\mathbf{X} = \{X_t \mid t = 0, 1, 2, \ldots, N\}$ and $\mathbf{Y} = \{Y_t \mid t = 0, 1, 2, \ldots, N\}$.

One can also study the correlation between the phases and amplitudes of the time series within normalized non-overlapping frequency bands, by first applying the Fourier transform to the time series. This tye of correlation is called "*spectral coherence*", and it is defined as

$$\mathrm{SpeCoh}(f) = \frac{|S_{XY}(f)|^2}{|S_{XX}(f)|\,|S_{YY}(f)|} \tag{2.32}$$

where $S_{XY}(f)$ is the cross-power-spectral density, and $S_{XX}(f)$ and $S_{YY}(f)$ are the auto-power-spectral densities for the variables w and s, respectively. Equation (2.32) defines a frequency-dependent correlation coefficient with values ranging between zero (no correlation) and unity (perfect correlation) for each frequency. The spectral coherence is widely used to find dynamic functional connectivity in brain networks [176] and in the study of climate teleconnections [177], among others.

The study of causal relationships in complex systems using correlation is contested for a number of reasons [161, 178, 179]. One of the main point is the presence of confounding factors, i.e. two variables X and Y which share a common cause Z can be correlated just because of the confounding effect of Z, leading to the inference of spurious relationships with inflated correlation coefficient [180]. Correlation coefficients can be corrected by statistically removing the influence of confounding factors. The *partial* correlation coefficient is a well-known measure of association between two variables, with the effect of a set of controlling variables removed. It is defined as the Pearson correlation between the partial residual e_x and e_y computed after the estimation of the parameters w_x and w_y, which are used respectively to explain x using z ad y using z in a multivariate regression parameter. Specifically, the partial correlation coefficient is expressed by

$$\boldsymbol{r}_{XY|Z} = \frac{N \sum_{i=1}^{N} e_{x,i} e_{y,i}}{\sqrt{N \sum_{i=1}^{N} e_{x,i}^2} \sqrt{N \sum_{i=1}^{N} e_{y,i}^2}} \tag{2.33}$$

where

$$\begin{aligned} \mathbf{w}_x^* &= \arg\min_{\mathbf{w}} \left\{ \sum_{i=1}^{N} \left(x_i - \langle \mathbf{w}, \mathbf{z}_i \rangle \right)^2 \right\} \\ \mathbf{w}_y^* &= \arg\min_{\mathbf{w}} \left\{ \sum_{i=1}^{N} \left(y_i - \langle \mathbf{w}, \mathbf{z}_i \rangle \right)^2 \right\} \\ e_{x,i} &= x_i - \langle \mathbf{w}_x^*, \mathbf{z}_i \rangle \\ e_{y,i} &= y_i - \left\langle \mathbf{w}_y^*, \mathbf{z}_i \right\rangle \end{aligned} \tag{2.34}$$

The computation should also take into account the time lag when studying causal relations in time series data.

Although largely used in applications, there is evidence that the structures resulting from correlations not only have no meaning in terms of causality, but also the subsequent analysis on the structural features may be undermined (see e.g. [178]). Therefore, the use and interpretation of such measures of association should always be conducted carefully. More sophisticated causal analysis tools should be preferred whenever possible.

Granger Causality Arguably, the most successful definition of causal dependence in complex dynamical systems is the one given by the Nobel laureate C.W.J. Granger in [181], and it is based on the idea that the variable X causally influences Y if the knowledge of the past values of X enhances the ability to *predict* future values of Y. Let Ω_n be the knowledge available in the universe at all times $t \leq n$; In [175] Granger introduces the two following axioms:

- **Axiom 1**: The past and present may cause the future, but the future cannot cause the past.
- **Axiom 2**: Ω_n contains no redundant information, so that if some variable $\mathbf{Z}$ is functionally related to one or more other variables, in a deterministic fashion, then $\mathbf{Z}$ should be excluded from Ω_n.

Therefore, considering two time series $\mathbf{X} = \{X_t \mid t = 1, \ldots, n\}$ and $\mathbf{Y} = \{Y_t \mid t = 1, \ldots, n\}$ the so called "Granger causality" is expressed by

$$P\left(X_{n+1} \in A \mid \Omega_n\right) \neq P\left(X_{n+1} \in A \mid \Omega_n - \mathbf{Y}\right) \tag{2.35}$$

in other words, Y causes X if the probability that a future value of the series X being in some set A , is different depending on whether or not we know something about the event Y. From this theoretical definition, Granger develops an operational definition which centers on building a forecast model for the time series. Assuming that the time series for X and Y are stationary and linear we define two alternative autoregressive processes:

$$\begin{aligned} f_1(\mathbf{X}) &= \left\{ X_t = \sum_{i=1}^{n} a_i X_{t-i} + \varepsilon_i \right\} \\ f_2(\mathbf{X}) &= \left\{ X_t = \sum_{i=1}^{n} b_i X_{t-i} + \sum_{j=1}^{n} c_j Y_{t-j} + \upsilon_{ij} \right\} \end{aligned} \tag{2.36}$$

where a, b and c are the coefficient of the autoregressive models, ε and υ are uncorrelated Gaussian random noises. Therefore, if the uncertainty on model f_2 – where the influence of Y is considered – is smaller than the uncertainty in model f_1, then Y Granger causes X. In practice, a vector autoregressive model can be used to test the statistical significance of the association:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} A^i_{xx} & A^i_{xy} \\ A^i_{yx} & A^i_{yy} \end{pmatrix} \begin{pmatrix} X_{t-i} \\ Y_{t-i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{pmatrix} \tag{2.37}$$

where i is the lag, and the coefficient A^i of the model are estimated by the maximum-likelihood method. The causal relation between X and Y is assessed by testing $A^i_{xy} \neq 0$ through a Wald test on the coefficient.

The Granger causality has the advantage of providing a clear operative definition of causality that can be easily implemented. On the other hand, the method, rigorously valid just for linear, stationary systems, represent a challenge when used to study real-world complex systems [182]. In fact, Granger causality requires separability, meaning the casual variable is independent of the variable that it influences, that is typical of stochastic and linear systems (on this point, see also the paragraph "Convergent Cross Mapping" on page 35). Variants on this method have been proposed to enhance the estimation accuracy of the autoregressive model parameters and to take into account the influence of possible latent confounding variables (see e.g. [183, 184]).

Mutual Information The methods described so far can only measure linear associations, but they would miss nonlinear ones. The mutual information is an information-theoretic measure that quantifies the amount of Shannon information that one variable convey about the another random variable, and can be applied also to non-linear systems. Let $\{X_t\}$ and $\{Y_t\}$ be two time series that are the temporal realizations (stochastic process) of two random variables with distributions $P(X_t = x_t)$ and $P(Y_t = y_t)$ respectively. Then, we define the *joint entropy* as:

$$\mathrm{H}(X_t, Y_t) = -\sum_t P(x_t, y_t) \log_2 P(x_t, y_t) \tag{2.38}$$

where $P(x_t, y_t) = P(X_t = x_t, Y_t = y_t)$ is the joint probability distribution of the two random variables. This quantity expresses the average information of the joint event (X, Y). If the two random variables are independent, then $P(x_t, y_t) \triangleq P(x_t) \cdot P(y_t)$ and consequently $\mathrm{H}(x_t, y_t) = \mathrm{H}(x_t) + \mathrm{H}(y_t)$ – from the definition of Shannon entropy. Thus, it follows that for any pair of

$$\begin{aligned} \mathbf{MI}(X_t, Y_t) &= \mathrm{H}(X_t) + \mathrm{H}(Y_t) - \mathrm{H}(X_t, Y_t) = \\ &= -\sum_t P(x_t, y_t) \log_2 \frac{P(x_t, y_t)}{P(x_t)\, P(y_t)} \end{aligned} \tag{2.39}$$

which is the definition of mutual information. This is equivalent to the Kullback-Leibler divergence that would be obtained in comparing the two processes X_t and Y_t as if they were independent with the actual empirical (joint) distribution of the data. The mutual information quantifies the "cost" – in terms of information – for encoding the pair (X_t, Y_t) as a pair of independent random variables, when in reality they are not. When mutual information equals zero, the two variables are independent. Conversely, a mutual information greater than zero indicate the strength of the association between the variables, regardless of the linearity.

The numerical estimation of the mutual information requires the empirical joint probability distribution $P(X_t, Y_t)$ to be computed, which is a non-trivial

task. Example of numerical computation of the mutual information can be found e.g. in [185].

It is to be noticed that the mutual information is a symmetric measure, implying that $\mathbf{MI}(X_t, Y_t) = \mathbf{MI}(Y_t, X_t)$, the direction of the causal relation, hence, cannot be derived applying the definition as in eq. (2.39). In addition, the *time* component has no role in the definition, neglecting the basic assumption that causes precede effects, as already mentioned. Introducing a time delay in one of the two variables, besides being partly arbitrary, still does not allow to take into account the influence of a third confounding variable that may act on both the variables understudy. Extensions to the method have been proposed, with applications in the study of complex systems of different kinds (see e.g. [186, 187]).

Transfer Entropy The temporal component inherent in the definition of cause-effect relation is restored in the definition of the "transfer entropy", that is an information-theoretic measure that quantifies the statistical coherence between systems evolving in time [188]. This measure, allows also a third set of variables to be included in the analysis, in order to explicitly remove their influence. The transfer entropy works under the assumption that the processes that we are observing are *k-Markovian*: let $\{X_t\}$ and $\{Y_t\}$ be two time series observed from stochastic processes with distributions $P(X_t = x_t)$ and $P(Y_t = y_t)$ respectively, and with joint distribution $P(X_t = x_t, Y_t = y_t)$. If the processes are stationary and Markovian of order k, then

$$\begin{aligned} P(x_{t+1} \mid x_t, \ldots, x_{t-k}, \ldots) &= P(x_{t+1} \mid x_t, \ldots, x_{t-k}) \\ P(y_{t+1} \mid y_t, \ldots, y_{t-k}, \ldots) &= P(y_{t+1} \mid y_t, \ldots, y_{t-k}) \end{aligned} \tag{2.40}$$

Therefore, the current value of two variables depends only on their previous k values. Let also

$$\begin{aligned} x_t^{(k)} &= \{x_t, \ldots, x_{t-k+1}\} \\ y_t^{(l)} &= \{y_t, \ldots, y_{t-l+1}\} \end{aligned} \tag{2.41}$$

then, the average number of bits needed to encode the observation at time $t+1$, once the previous k and l values took place, is given by the conditional Shannon entropy

$$H\left(x_{t+1} \mid x_t^{(k)}\right) = -\sum_t P\left(x_{t+1}, x_t^{(k)}\right) \log_2 P\left(x_{t+1} \mid x_t^{(k)}\right) \tag{2.42}$$

$$H\left(y_{t+1} \mid y_t^{(l)}\right) = -\sum_t P\left(y_{t+1}, y_t^{(l)}\right) \log_2 P\left(y_{t+1} \mid y_t^{(l)}\right) \tag{2.43}$$

In the same way, we can compute the average information about the future value x_{t+1} knowing its history $x_t^{(k)}$ and also the past values of the other variable $y_t^{(l)}$.

This is given by

$$H\left(x_{t+1} \mid x_t^{(k)}, y_t^{(l)}\right) = -\sum_t P\left(x_{t+1}, x_t^{(k)}, y_t^{(l)}\right) \log_2 P\left(x_{t+1} \mid x_t^{(k)}, y_t^{(l)}\right) \tag{2.44}$$

Equipped with these expressions, we can now ask *how much information would we loose if we assume that the process generating X does depend just on its past, and does not depend on the past of Y?* The answer is given by the difference in the expected information in the two cases, with and without the information on Y, that is, by computing the Kullback-Leibler divergence for those two situations:

$$\mathbf{T}_{Y\to X}^{(k,l)} = \sum p\left(x_{t+1}, x_t^{(k)}, y_t^{(l)}\right) \log \frac{p\left(x_{t+1} \mid x_t^{(k)}, y_t^{(l)}\right)}{p\left(x_{t+1} \mid x_t^{(k)}\right)} \tag{2.45}$$

that corresponds to the difference of eq. (2.42) and eq. (2.44). Equation (2.45) is the definition of the *transfer entropy*, which is a measure of the information about the past of Y "flowing" into the future value of X. Note that this is not just a matter of statistical association, as for the methods previously presented, since eq. (2.45) encompasses a model for the the *dynamics* of the system, that is assumed to be Markovian, and hence includes the temporal component of the information flow. The transfer entropy is equal to zero if and only if X conditional on its own past, is independent of the past of Y. In addition, any pair of time series will have two transfer entropies, the first is the one in eq. (2.45), the second is given by flipping the order of X and Y. This makes possible the comparison to determine also the verse of the causal relation.

The transfer entropy provides an information-theoretic definition of causality as the presence of an "information flow" between two events. It can handle both linear and non-linear systems and unlike mutual information, it ignores static correlations due to the common history or common input signals [188]. Interestingly, the transfer entropy is equivalent to the Granger causality when the variables are Gaussian [189]. For this method to work, the two time lag k and l must be fixed. In fact, eq. (2.41) defines a state dependency of x and y on their k and l past observations, analogous to the state space reconstructed by delay-embedding following the Takens' theorem [190]. For an overview of the methods to choose the lag and numerically compute the transfer entropy see e.g. [191] and references therein. It is to be noticed, that information transfer and causality are strictly related but distinct concepts. We refer the reader to [192] for further discussion on this relationship.

Convergent Cross-Mapping The convergent cross mapping (CCM) is a methodological approach, developed by Sugihara et al. in [47], used to detect causal relationships in complex dynamical systems. In particular, the CCM assumes that a

causal relationship between two variables x and y exists, if their dynamics share an underlying common attractor manifold and the state of the causal variable can be used to determine the state of the other. Causality is thus interpreted as a coupling in the dynamics of the two variables x and y.

The mathematical foundation of the method is the Takens' theorem, which provides the conditions under which the topology of an attractor can be reconstructed from a vector of observations. If the dynamical system is governed by the following system of differential equation

$$\dot{x} = F\left(x\left(t\right)\right) \tag{2.46}$$

where $F : \mathbb{R}^N \to \mathbb{R}^N$ is a generic function of the system state. We also assume that the state is located on a manifold $\mathcal{M} \subset \mathbb{R}^N$. The observer only see the system state through a measurement function

$$y(t) = f\left(x\left(t\right)\right).$$

The evolution of the state variable can be described in terms of its initial condition by using a function $G : \mathcal{M} \times \mathbb{R} \to \mathcal{M}$ such that

$$x\left(t_0 + \tau\right) = S_\tau\left(x\left(t_0\right)\right)$$

where τ is equal to the sampling time of the observations. Therefore, for any $k \in \mathbb{N}$,

$$x\left(t_0 + k\tau\right) = S_\tau^{\{k\}}\left(x\left(t_0\right)\right) = S_\tau^1 \circ S_\tau^1 \circ \cdots \circ S_\tau^k\left(x\left(t_0\right)\right)$$

We define the E-delay coordinate map as

$$\begin{aligned} D(x(t)) &= \left[y(t), y\left(t-\tau\right), \cdots, y\left(t-(E-1)\tau\right)\right]^T \\ &= \left[f(x(t)), f \circ S_{-\tau}(x(t)), \cdots, f \circ S_{-\tau}^{E-1}(x(t))\right]^T. \end{aligned} \tag{2.47}$$

The Takens theorem specifies the conditions under which the function D is an embedding of the unobserved manifold $\mathcal{M}$ in the *reconstruction space* $\mathbb{R}^N$. More explicitely, Takens in [190] gave the following

Theorem (Takens embedding theorem). *Let $\mathcal{M}$ be a compact manifold of dimension μ and that the dynamics defined in eq. (2.46) is confined on $\mathcal{M}$. If the periodic points of $S_\tau^{\{k\}}$ with periods $\pi \leq 2\mu + 1$ are finite in number, and the eigenvalues of $S_\tau^{\{k\}}$ are different and different from 1, then for pairs (f, S) such that $f : \mathcal{M} \to \mathcal{M}$ is a smooth diffeomorphism and $S : \mathcal{M} \to \mathbb{R}^N$ is a smooth function, it is a generic property that the map $D\left(x\left(t\right)\right) : \mathcal{M} \to \mathbb{R}^{2\mu+1}$*

$$D(x(t)) = \left[f(x(t)), f \circ S_{-\tau}(x(t)), \cdots, f \circ S_{-\tau}^{2\mu}(x(t))\right]^T \tag{2.48}$$

is an embedding.

The convergent cross mapping relies on this theorem to reconstruct a "shadow" manifold $\tilde{M}_x$ starting from a single time series $x(t)$, corresponding to the causal variable, since each variable contains information about all the others. Thanks to the Takens theorem we know that the shadow manifold is diffeomorphic to the true manifold. This manifold is reconstructed through the delay coordinate map of eq. (2.47), where the values for E and τ are determined by using one of the many dedicated algorithms[4] (see e.g. [195, 196, 197, 198, 199]). The validity of the reconstruction of $\tilde{M}_x$ can be assessed by simulating the dynamics of x_t from the manifold and comparing the result with the measured values of x_t. The manifold is correctly reconstructed if the out-of-sample predictability – measured with error statistics – is significant [47].

In the same way, a second shadow manifold $\tilde{M}_y$ for the variable y is reconstructed and the points with a "similar history" on $\tilde{M}_x$ are used to estimate the values on the other manifold $\tilde{M}_y$. The CCM determines how well the points on $\tilde{M}_x$ corresponds to the points on $\tilde{M}_y$. The "similar history" is defined by taking a minimum bounding simplex around a point $\tilde{x}_t$. The minimum number of points to define a bounding simplex in an E-dimensional space is $E+1$, therefore, the $E+1$ nearest neighbours $\tilde{x}_{t'_i}$ of $\tilde{x}_t$ are identified, and the Euclidean distances between each neighbour and $\tilde{x}_t$ are computed as $d_i = \left|\tilde{x}_t - \tilde{x}_{t'_i}\right|$ with $i = 1, ..., E+1$. To each neighbour is then associated a weight w_i defined as

$$w_i = \frac{e^{\frac{-d_i}{d_{min}}}}{N} \tag{2.49}$$

where d_min is the minimum distance between the point $\tilde{x}_t$ and every neighbour $\tilde{x}_{t'_i}$, and $N = \sum_{i=1}^{E+1} e^{\frac{-d_i}{d_{min}}}$. The weights are used to estimate the current value of the variable y. The conditional estimated value of y is given by

$$\hat{y}_t \mid \tilde{x}_t = \sum_{i=1}^{E+1} w_i t_{t'_i} \tag{2.50}$$

Since the shadow manifold $\tilde{M}_x$ is diffeomorphic to M, the *cross-mapping* estimate $\bar{y}_t$ will *converge* to the real value y as the length of the time series ("library length") goes to infinity[5]. This procedure is repeated for each point on the shadow manifold and also switching x and y to obtain $x \mid \tilde{y}_t$.

[4] According to the Whitney theorem, the diffeomorphism from the real to the shadow manifold is ensured by choosing $E \geq 2\mu + 1$ [193] and the result may be generalized also to manifold with fractal dimension, such as the strange attractors [194, 195]

[5] The values of $\bar{y}_t$ converge to the measured y_t only in purely deterministic systems. In real-world application, the shadow manifold is an approximated reconstruction of the real manifold, and the convergence is limited by the observational errors and process noise.

Finally, the Pearson correlation between the real value of y and its estimate through the shadow manifold $y_t \mid \tilde{x}_t$ is computed as

$$\mathcal{C}_{yx} = \rho\left(y, y_t \mid \tilde{x}_t\right) \tag{2.51}$$

If the variable y act as a cause for the variable x, the correlation $\mathcal{C}_{yx}$ will grow to significant levels as the library length increase. Moreover, the CCM can be used to determine also the direction of the causal relations by computing the correlation difference[6]

$$\Delta_{\mathcal{C}} = \mathcal{C}_{yx} - \mathcal{C}_{xy} = \rho\left(y, y_t \mid \tilde{x}_t\right) - \rho\left(x, x_t \mid \tilde{y}_t\right). \tag{2.52}$$

A value of $\Delta < 0$ means that (at a fixed library length) cross mapping x using y gives better estimation than cross mapping y using x. Therefore, x is a driver of y. If the two variables are mutually coupled the causal relation is bidirectional and the estimate $\hat{y}_t \mid \tilde{x}_t$ should converge to the observed time series y_t as well as the estimate $\hat{x}_t \mid \tilde{y}_t$ should converge to x_t. The CCM recognizes also the transitivity in causal networks, i.e. those causal structure in which two variables x and y (that may or not interact) are influenced by a common variable z. The true causal relations can be distinguished from spurious correlations due to the confounding factor z [47].

Unlike Granger causality, the CCM can detect the right causal structure even in non-separable systems, where feedbacks and non-linearity are present and the system cannot be considered part by part but it must be understood as a whole. In addition to separability, the CCM differs from the Granger causality and also from transfer entropy, because it does not aim at predicting the future behaviour of a variable, but rather it estimates the *correspondence* of states of a variable with those of another, on a common attractor manifold.

For this method to work properly, are required long time series since the convergent occurs with large library length that determines sufficiently "dense" reconstruction of the shadow manifold. There are some extensions of the method outlined here, that address this and other issues of CCM (see e.g. [200, 201, 202]).

2.3 Novel perspectives

In the following chapters we will use the methods described in the previous sections to study various aspects of network modelling and human mobility. In the next chapter, we will present our novel framework for the study of the topological features of complex networks when the structure of the system is uncertain. The

[6]Note that the value of correlation and consequently the value of Δ is also a function of the library length, and the causal relation should be identified looking at the trend of these statistics with growing library length.

novelty lies in the probabilistic description of the complex system structure, that is considered as a "fuzzy network" where the causal/functional interactions might be present with some probability that depends on the "strength" of the interactions measured through the methods presented in the previous sections. The topological features, such as the degree and the clustering, are consequently redefined as probability distributions rather than punctual values, convening additional information on the uncertainty of the network structure.

The mathematical methods for causal analysis are applied also in the chapter 4 in which we derive the relation between a large set of variables, including human mobility and environmental conditions. We will introduce a new information-theoretic method to detect abrupt shifts in the dynamics of a system and we offer a comprehensive insight on the effects of a sudden change in human activities on the surrounding environmental conditions.

The gravity model introduced in section 2.1 is used in chapter 5 to understand which factors are important in determining the mobility flows during a catastrophic environmental event. The gravity model is estimated using a regularized ridge regression.

In the chapter 6, we present our generalized version of the radiation model (shortly introduced in section 2.1) which allows any type of feature to be considered as a driver of migration, besides the population, without altering the physical process of the original model. Our Feature-Enriched Radiation Model can be used when location features are deemed important for the prediction of mobility fluxes.

The modelling of large-scale human mobility patterns is made difficult because of the lack of reliable and consistent data. The existing dataset on large-scale mobility flows and migration are often incomplete and/or incomparable because of the different definitions of a migrant, non-migrant and commuters persons given by national and international agencies and because of the different methods used to estimate bilateral fluxes. This fact has had an unforeseeable impact on the progress of this thesis, that had to be oriented towards the advancement of the methodology in both complex network science and human mobility. In the following of this thesis, we do not provide the conclusive argument or solution for the prediction of human mobility and migration patterns, but rather, we will illustrate the mathematical and physical issues that arise in studying the complex human mobility system and offer possible solutions to them. Specifically, we will illustrate what methodological and applied contributions we made, and also what are the future research directions that may help framing the problem the right way, within a "holistic" (not the new-age, but the Aristotelian) view solicited by researchers and institutions [10, 49, 203].

Chapter 3

The structure: measuring topological descriptors of complex networks under uncertainty

In brief

Revealing the structural features of a complex system from the observed collective dynamics is a fundamental problem in network science. In order to compute the topological descriptors that characterize the structure of a complex system (e.g. the degree, the clustering coefficient), it is usually necessary to reconstruct the underlying network by using the methods presented in section 2.2. In this setting, the uncertainty about the existence of the edges is reflected in the uncertainty about the topological descriptors. In this chapter we propose a novel methodological framework to evaluate this uncertainty, replacing the topological descriptors, even at the level of a single node, with appropriate probability distributions, eluding the reconstruction phase. Our results provide a grounded framework for the analysis and the interpretation of widely used topological descriptors in scenarios where the existence of network connectivity is statistically inferred or when the probabilities of existence π_{ij} of the edges are known. To this purpose we also provide a simple and mathematically grounded process to transform the discriminating statistics into the probabilities π_{ij}.
This chapter is published in *Physical Review E* [12].

3.1 Introduction

> As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.
>
> *Geometry and Experience* – 1921
> Albert Einstein

Complex natural and artificial systems are composed of many interacting dynamical units which exhibit a collective behavior [204]. This is the result of the interplay between the dynamics of the constituents and the interactions among them. The structure of the interactions and the (nonlinear) dynamics have to be considered simultaneously to model such systems [205]. Unfortunately, the structure of many empirical systems usually remains hidden, but the dynamics of some physical quantity can be observed and measured. From such observations the connectivity can be inferred for a broad class of systems [206, 207], from the human brain [208, 209, 210, 211], to financial [212, 213], weather and climate systems [214, 215], including hydrological processes [216, 217] and biological systems [218, 219]. As explained in section 1.2, human mobility patterns emerge from the relations between various systems and factors, from economic to social and environmental conditions. The relations between these systems are often either hidden or uncertain, therefore the connectivity should be characterized taking into account the lack of information contained in the data. In this chapter we present a new way to model the *structure* of a complex system, that allows the uncertainty to be included in various topological descriptors. Borrowing the vocabulary of neuroscience, we can distinguish different types of connectivity: structural, effective and functional connectivity [220]. The structural connectivity refers to the existence of some physical relations connecting different parts of a complex system (such as the synapses in brain, or the roads in a urban system). The functional connectivity instead, refers to the statistical dependence of the signals coming from different parts of the complex system (such as the variation in income in different regions or the changing environmental conditions). Some researchers, especially in network neuroscience, also talk about effective connectivity, which brings in the element of causation: two variables are connected if one is the cause of the behaviour of the other. Functional and effective connectivity are usually inferred using the methods described in section 2.2 and represent the map to understand the dynamical behavior of the system, to figure out which parts are mostly affected by perturbations in others, and so on. In practical applications, the functional and effective connectivity are often considered as a proxy for the structure. It is worth noting that the distinction between these three types of connectivity is not always

completely clear nor it is explicit, for example in the climate teleconnections [221].

Many topological descriptors are used to characterize the structural features of a complex system (e.g. the degree, the transitivity, etc.), but to compute them an earlier reconstruction of the structure itself is usually necessary. The goal of network reconstruction is typically to solve this inverse problem [222]: from information about the dynamics, reconstruct the network of interactions. In general, a complex system can be described as follows: let $\mathbf{x_i}(t)$ denote the internal D-dimensional state $\mathbf{x_i}(t) = [x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(D)}]^{\mathrm{T}}$ of a system consisting of N dynamical units, at time t. The evolution of the state is governed by the system of N ordinary differential equations

$$\dot{\mathbf{x}}_i(t) = \Psi_i(\mathbf{x_i}(t), \boldsymbol{\gamma}_i) + \sum_{j=1}^{N} A_{ij}\Phi_{ij}(\mathbf{x_i}, \mathbf{x_j}) + \mathbf{u_i}(t) + \boldsymbol{\eta_i}(t)$$

where $i, j \in \{1, 2, \ldots, N\}$, $t \in \mathbb{R}$; the function $\Psi_i : \mathbb{R}^D \to \mathbb{R}^D$ and $\Phi_{ij} : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$ respectively define the intrinsic and interaction dynamics of the D-dimensional units. The function $\mathbf{u}(t)$ represents external drivers, $\boldsymbol{\eta}(t)$ is a dynamic noise term and $\boldsymbol{\gamma_i}$ is a set of dynamic parameters. Finally, the term A_{ij} defines the interaction topology in terms of the adjacency matrix $\mathbf{A}$ such that $A_{ij} = 1$ if there is a direct physical interaction from unit j to i and $A_{ij} = 0$ otherwise. This matrix completely defines a network, that is, an abstraction used to model a system that contains discrete, interconnected elements. The elements are represented by nodes (also called vertices) and the interconnections are represented by edges. In general, one should take into account the response of the experimental setup used for measuring the state (and the measurement noise), resulting in a vector $s(\mathbf{x}(\mathbf{t}))$ of measured observables which is a function of $\mathbf{x}(t)$. In many cases, the reconstruction problem relies solely on the vector $s(\mathbf{x}(\mathbf{t}))$, a multivariate time series. Many different methods have been proposed to recover the structure of the interactions between dynamical units from time series (see the section 2.2 and also [223, 224, 225, 226, 227, 228, 202, 229, 212, 230, 231]). The methods presented in section 2.2 consist in the quantification of the interactions between units through an appropriate discriminating statistic (pairwise correlations [232], statistical causality between units [233], mapping information flow from the observed collective dynamics [234]) and then to apply a criterion to decide whether the measured interaction is significant or not [235, 236, 237]. The choice of the criterion is crucial, but typically it introduces some arbitrary choices in the process. The current reconstruction procedures often rely on heuristics to choose a threshold value for the pairwise correlation or causality measures. Values below the threshold are discarded, so that an edge is assigned only between units whose interaction is sufficiently strong. This procedure is known to produce complex features even when no complex structure is present [238]. Preferably, using a more sophisticated

statistical analysis, a set of *p-values* is computed to evaluate the significance of the edge between the nodes with respect to a null-model [239]. However, even in this case, the process incurs in the issues of partial correlations [225] and multiple testing [240]. Other approaches have been recently proposed (e.g. [241, 242]), by which the posterior probability distribution of the network structure is computed using suitable generative processes and prior information. The network is reconstructed by sampling from this distribution. These approaches require the model of the dynamics to be defined together with its corresponding probabilistic model for the data. In other cases, the network structure can be constructed from static observations and ad-hoc measurements (see e.g. [243, 244, 245]), which may be affected by noise and measurement error. In any case, after the reconstruction of the network structure, it is possible to compute the topological descriptors of the structure.

In this chapter, we propose a new methodological framework to analyse the structural features of a complex network when its topological connectivity is specified by edge probabilities, without the explicit reconstruction of the network structure. In addition, we propose a simple procedure to obtain the edge probabilities, given the *p-values* that quantify the supporting evidence of the related discriminating statistics. The network descriptors are redefined as stochastic variables, whose probability distributions can be used to infer the relevant statistics and to evaluate their robustness against the uncertainty. Note that this framework complements other approaches like [246] – which regards community detection – and [241, 242] – where a specific generative model for the data is used. In fact, our approach aims at computing the topological descriptors of a complex network having information about the edge existence, without reconstructing the entire network structure and in the absence of a model for the dynamical process. Moreover, the proposed method does not build on assumptions about the topological features of the underlying network, nor on its generative process, but rather includes the prior knowledge about the existence of each edge. Employing a Bayesian procedure we derive for every i and j the probability π_{ij} that the node i is linked to the node j, given the *p-value* from the above mentioned analyses. Hence, the actual complex network is considered as a realization from the possibilities encoded in the probabilistic model that we call "fuzzy network" model. Under this probabilistic perspective, all the network descriptors must be redefined as random variables. A natural way to recover the descriptive information is to consider the whole distribution or a suitable statistic. Therefore, we have defined the "fuzzy" counterpart of some basic structural descriptors such as the node degree and the network expected degree, the clustering coefficient, and the probability of having a unique connected component. For each of them, we present the analytical probability distributions and the main statistics. We applied this framework to various well-

known synthetic and real-world networks starting from multivariate time series, and compared the results to the ones from a classical reconstruction method.

3.2 Analysis of network connectivity under uncertainty

This section describes in detail the process to derive the network descriptors from the observed multivariate time series through the fuzzy network model. Given the time series of the nodes' dynamics, a pairwise connectivity measure is computed for each pair of nodes. Subsequently, a bootstrap method is performed for each pair of nodes to derive a *p-value* for the connectivity. Furthermore, relying on the Bayes theorem, the *p-values* are translated into the posterior probabilities of existence of the edges. Consequently, the probabilities are used to define the fuzzy network and the stochastic network descriptors. It is worth remarking that this is only a specific way to obtain a probability for each connection in the system: other approaches, based for instance on inference with explicit generative models [241, 247] can be used. In fact, the following analysis does not depend on the specific method to obtain probabilities, which are used as input parameters, so that a wider set of problems can be addressed, in which the probabilities π_{ij} are directly provided instead of time series. Nevertheless, the study of complex time-varying dynamical systems through time series and connectivity measures has a great explanatory power and significant practical importance. Hence we focused our discussion and numerical experiments on this type of systems. For the sake of simplicity, in what follows we assume the networks to be undirected and unweighted.

3.2.1 Connectivity Matrix

The procedures commonly adopted to reconstruct the network topology of a complex system rely on some statistical descriptor used as a proxy for the structural connectivity of the system. These descriptors are able to quantify the relationship between the dynamics of the system's components. In the following, we will apply three types of statistical relation: the Pearson correlation coefficient (CC), the Spearman's rank correlation (SC) and the Spectral Coherence (SpeCoh) [248]; an information-theoretic tool: the mutual information (MI) [249]; and a state-space reconstruction tool, namely the Convergent Cross Mapping (CCM) [47]. These methods have been applied to reconstruct complex networks in different contexts, from neuroscience [250, 251, 252, 253] to climatology [254, 217, 255, 256], finance [257, 213] and ecology [47]. As extensively reported in section 2.2, the problem is to quantify the evidence of the interaction between two components using the information enclosed in the time course of the state vector. The analysis is conducted

pairwise, for each pair of components. The result is a matrix which summarizes the strength of the interaction between each pair. We call this matrix "connectivity matrix" $\mathbf{C}$ to distinguish it from the adjacency matrix

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & \dots & c_{1,N} \\ \vdots & \ddots & \vdots \\ c_{N,1} & \dots & c_{N,N} \end{bmatrix}. \tag{3.1}$$

3.2.2 Probabilities of existence

The statistical significance of the values in the connectivity matrix can be quantified deriving the corresponding *p-values*. To do so, we perform a surrogate data analysis using the reshuffled version of the time series to compute a null model (see section 3.3 for more details), which expresses the null hypothesis H^0_{ij} of lack of connectivity between nodes i and j. The result of this process is a matrix of p-values p_{ij} which quantifies – for each possible edge e_{ij} – the strength of the evidence against the null hypothesis H^0_{ij}. In the usual reconstruction context the p_{ij} can be used (after adjusting them for the multiplicity) to test against the null hypothesis of lack of connectivity [239]. These would lead directly to the reconstructed adjacency matrix of the network given a level of significance fixed a priori.

Instead, we ask for the *probability* that the null hypothesis H^0_{ij} is true, given the p_{ij}. That is equivalent to asking for the probability of existence of the edge e_{ij} once the corresponding *p-value* is known, which reads

$$P(H^1_{ij}|p_{ij}) = 1 - P(H^0_{ij}|p_{ij}) = \pi_{ij}. \tag{3.2}$$

To derive this probability we rely on the work of [258] and [259], which provide a Bayesian argument to obtain the posterior probability distribution $P(H^0_{ij}|p_{ij})$ for the null hypothesis H^0_{ij} given the *p-value* (on the rhs of section 3.2.2). To determine the functional form of $P(H^0_{ij}|p_{ij})$ from the Bayes theorem, the distribution of the *p-values* under the null and alternative hypotheses are needed. It is known that the *p-values* under H^0_{ij} are distributed uniformly like Unif$(0, 1)$. This is a direct consequence of the Probability Integral Transform applied to the *p-values* [260]. Instead, under the alternative hypothesis H^1_{ij} the p_{ij} can be considered distributed as a Beta$(\xi, 1)$ probability distribution. This choice reflects the fact that the p_{ij} are bounded between 0 and 1 and that under the alternative hypothesis they are skewed on the left (toward 0). Since the standard Uniform distribution is a particular case of the Beta distribution ($\xi = 1$), it follows that the distribution of p_{ij} is

$$p_{ij} \sim f(p_{ij}|\xi) = \xi p_{ij}^{\xi-1}$$

so that the parameter ξ includes the information on which hypothesis is considered. In the Bayesian framework, given a prior distribution $g(\xi)$ for the parameter ξ, the test of the null against the alternative hypotheses is assessed by the Bayes factor

$$B_g(p_{ij}) = \frac{\mathrm{P}\left(p_{ij}|H_{ij}^0\right)}{\mathrm{P}\left(p_{ij}|H_{ij}^1\right)} = \frac{f(p_{ij}|1)}{\int_0^1 f(p_{ij}|\xi)g(\xi)d\xi} \tag{3.3}$$

By using the First Mean Value theorem and after some calculations, the inferior Bayes factor is obtained as

$$\begin{aligned} B_{ij} &= \inf_\xi B_g(p_{ij}) = \frac{f(p_{ij}|1)}{\sup_\xi \xi p_{ij}^{\xi-1}} \\ &= -ep_{ij}\log p_{ij} \quad \text{for} \quad p_{ij} < e^{-1} \end{aligned} \tag{3.4}$$

and $B_{ij} = 1$ for $p_{ij} > e^{-1}$ where e is the Euler's number. B_{ij} is independent on the parameter ξ and it is valid for any prior distribution on ξ. This can be interpreted as a lower bound for the odds of H_{ij}^0 on H_{ij}^1 given the form of the distribution under H_{ij}^1 [258]. Finally, using the definition, the (inferior) Bayes factor can be mapped into the minimum posterior probability for the null hypothesis given the *p-value* :

$$1 - \pi_{ij} = \left(1 + \left(\frac{B_{ij} \cdot P(H_{ij}^0)}{1 - P(H_{ij}^0)}\right)^{-1}\right)^{-1} \tag{3.5}$$

This formula gives the (maximum) posterior probability π_{ij} that the edge e_{ij} exists given the *p-value* from its connectivity measure, where $P(H_{ij}^0)$ is the prior probability for the null hypothesis, which is the only parameter to be fixed in this procedure. This parameter contains the prior knowledge about the possibility of finding an edge between two nodes. In principle, it can assume a different value for every edge in the network, depending on the amount of prior information available at the edge-specific level. In situations where a local characterization of the structure is unavailable, the $P(H_{ij}^0)$ can be unique and equal for all the edges, so that $P(H_{ij}^0) = P(H^0)$. For instance, a global value can be determined considering information about other networks (e.g. using the expected density of a set of known networks similar to the one under study) or with other problem-specific knowledge; otherwise, an uninformative prior can be used.

3.2.3 Building the Fuzzy Network

The probabilities π_{ij} of existence of the edge between nodes i and j can be rearranged in a matrix $\mathbf{P}$, to obtain the probabilistic counterpart of the adjacency

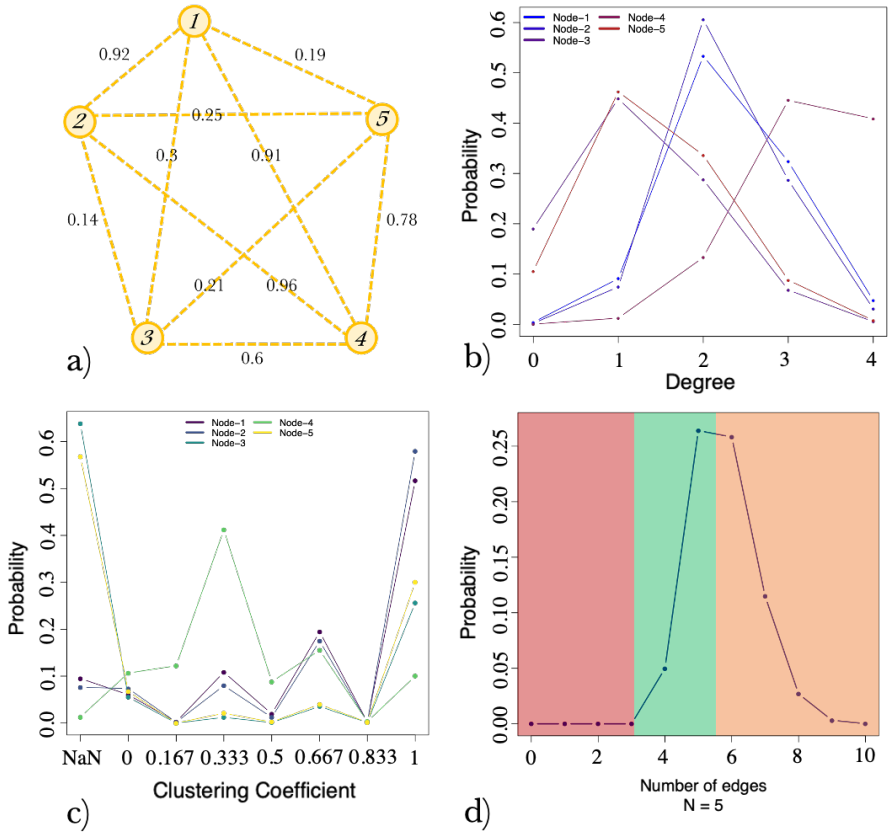


Figure 3.1 – **Probability distributions** for the degree (b) and the local clustering coefficient (c) for the toy network in (a). The figure (d) represents the probability of having a network of five nodes consisting of a single totally connected component with i edges (x-axis).

matrix:

$$\mathbf{P} = \begin{bmatrix} \pi_{1,1} & \dots & \pi_{1,N} \\ \vdots & \ddots & \vdots \\ \pi_{N,1} & \dots & \pi_{N,N} \end{bmatrix} \tag{3.6}$$

The matrix $\mathbf{P}$ resembles a weighted adjacency matrix, but it has a different meaning: the value π_{ij} is not a weight, but it represents the probability of existence of the corresponding edge. Therefore, the matrix $\mathbf{P}$ totally defines a complete network, whose edges might exist with a certain probability (see fig. 3.1). This representation, encodes all the knowledge about the structural connectivity of the network. We name this model "fuzzy network".

Given the stochastic nature of the edges, all the structural descriptors must be redefined as random variables. In what follows, we redefine some of the most widely used structural descriptors on the basis of the fuzzy network model.

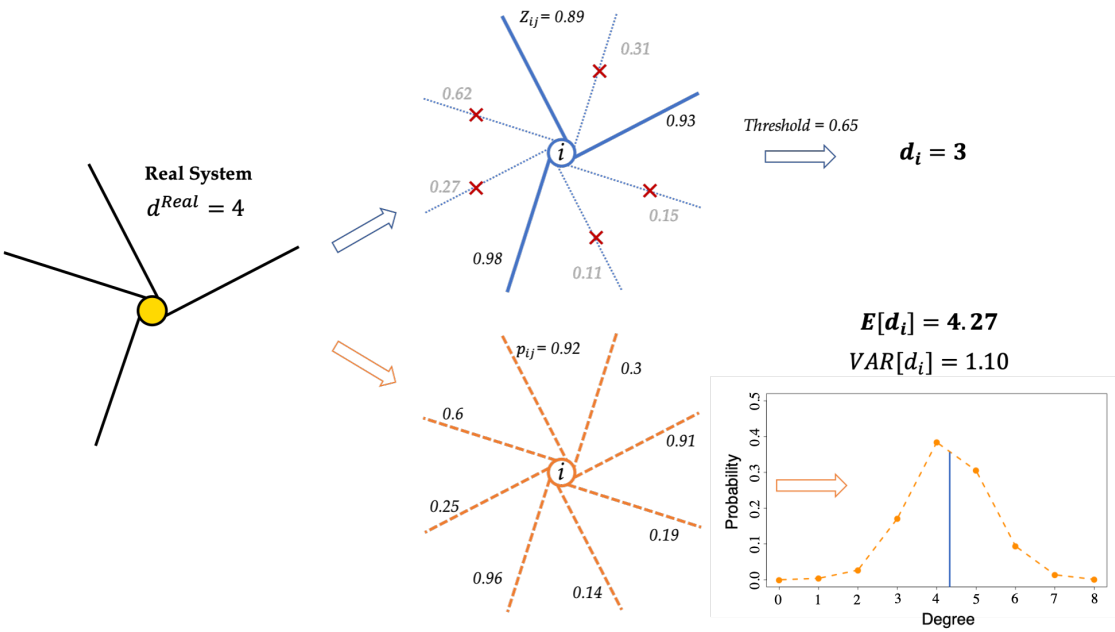


Figure 3.2 – **Connectivity reconstruction** of a toy system consisting of four edges. Comparison between the widely used thresholding technique (top, blue color), and the redefinition of the node degree as a random variable (bottom, orange color), for an hypothetical real node with degree 4. The node is represented as a node of a "fuzzy network" (below) in which a probability of existence is associated to each edge. The node degree distribution is plotted on the right-hand side, along with its mean and variance. The expected value results to be closer to the real value than the value of the thresholding process.

3.2.4 Node degree

Let us consider a single node in the fuzzy representation of the complex network (see fig. 3.2). The node i has N edges incident to it, each with an associated independent probability of being present. Under this condition, the usual definition of the node degree (i.e. the number of edges incident to the node) is no more applicable, since the node has all the possible degrees at the same time, each with a certain probability. Therefore, another definition of the degree is needed to take into account the uncertainty about the existence of the edges. The most natural choice is to define the degree as a random variable described by its probability distribution, which depends on the probabilities π_{ij}. The probability that the node i has degree $d_i = k$ can be thought of as the probability to have k successes in a sequence of N independent Bernoulli trials with success probabilities $p_{i1}, p_{i2}, ..., p_{i(N)}$:

$$e_{ij}|\pi_{ij} \sim Bernoulli(\pi_{ij}) \quad , \quad d_i = \sum_{j=1}^{N} [e_{ij}|\pi_{ij}] \tag{3.7}$$

If the π_{ij} were all equal, the probability distribution of the latter sum would be the well-known Binomial distribution. But in this case all the edges incident to node i have different probabilities of existence. Consequently, the probability of having k successful trials out of a total of N can be written as

$$P(d_i = k) = \sum_{\Lambda \in F_k} \prod_{j \in \Lambda} \pi_{ij} \prod_{l \in \Lambda^c} (1 - \pi_{il}) \tag{3.8}$$

where F_k is the set of all subsets of k edges that can be selected from $\{e_{i,1}, e_{i,2}, e_{i,3}, \ldots, e_{i,N}\}$. For example, if $N = 3$, then

$$F_2 = \{\{e_{i,1}, e_{i,2}\}, \{e_{i,1}, e_{i,3}\}, \{e_{i,2}, e_{i,3}\}\}.$$

Λ^c is the complement of Λ, i.e. the set $\Lambda^c = \{e_{i,1}, e_{i,2}, e_{i,3}, ..., e_{i,N}\} \setminus \Lambda$. This distribution is the so called Poisson-Binomial distribution, and it represents the node degree distribution. As for the Binomial, the mean is equal to the sum of the π_{ij} and the variance is the sum of the probabilities of success times the probabilities of fail:

$$\mu_{d_i} = \sum_j \pi_{ij} \quad , \quad \sigma^2_{d_i} = \sum_j \pi_{ij}(1 - \pi_{ij}) \tag{3.9}$$

Having an entire distribution for each node, we are provided with more information with respect to the case of the usual degree. This additional information makes the calculation of the degree more robust against uncertainty, since it is possible to compute the most significant moments of the distribution.

3.2.5 Expected degree of a network

An important summary quantity which characterizes a network is the expected degree of the network. In order to find the expected value for the entire network we exploit the properties of the Poisson-Binomial distribution in eq. (3.8). First of all, the Poisson-Binomial distribution is very well approximated by the Normal distribution for fairly small samples (the approximation can also be refined using a continuity correction for discrete random variables). This is a consequence of the Central Limit Theorem (CLT). More precisely, since the Poisson-Binomial is defined as the sum of independent but not identically distributed Bernoulli variables (see eq. (3.7)), the CLT needs to be considered in the Lyapunov formulation, which imposes a condition on the moments of the distribution of e_{ij} in eq. (3.7) [261]. Suppose $\{d_1, d_2, ...d_n\}$ is a sequence of independent random variables, each with finite expected value μ_{d_i} and variance σ_i. Let's define $s^2_n = \sum_{i=1}^n \sigma^2_{d_i}$. If for some $\delta > 0$, the Lyapunov's condition

$$\frac{1}{s_n^{2+\delta}} \sum_{i=1}^{n} \mathbb{E}\left[|d_i - \mu_{d_i}|^{2+\delta} \right] = 0 \tag{3.10}$$

is satisfied, then the sum of $\frac{d_i - \mu_{d_i}}{s_n}$ converges in distribution to a standard Normal random variable, as n goes to infinity:

$$\frac{1}{s_n} \sum_{i=1}^{n} (d_i - \mu_{d_i}) \xrightarrow{d} \text{Norm}(0, 1) \tag{3.11}$$

For the sum of Bernoulli random variables, the Lyapunov condition is easily satisfied (see the *Appendix B* for a detailed discussion) and the convergence is reached even for very small N; thus for a network having N nodes, the degree of node i follows:

$$d_i \sim PB\left(\pi_{i,N-1}, \ldots, \pi_{i,N-1}\right) \xrightarrow{d} \text{Norm}\left(\mu_{d_i}, \sigma_{d_i}\right)$$

where the parameter μ_{d_i} and σ_{d_i} are given by eq. (3.9). From the properties of the Normal distribution, and from eq. (3.9), it follows that

$$\sum_{i}^{N} d_i \approx \text{Norm}\left(\sum_{i}^{N} \mu_{d_i}, \sum_{i}^{N} \sigma_{d_i}\right) \tag{3.12}$$

In general, the total number of edges m in the network, and the expected degree c can be computed as

$$m = \frac{1}{2} \sum_{i}^{N} d_i \quad , \quad c = \frac{2m}{N} \tag{3.13}$$

which in this case are random variables, since the element d_i is stochastic. Consequently, we can compute the expected value of the random variable c taking into account section 3.2.5:

$$\begin{aligned} \mathbb{E}[c] = \mathbb{E}\left[\frac{2m}{N}\right] = \frac{1}{N}\mathbb{E}\left[\sum_{i}^{N} d_i\right] \approx \\ \frac{1}{N}\sum_{i}^{N} \mu_{d_i} = \frac{1}{N}\sum_{i}^{N}\sum_{j}^{N} \pi_{ij} \end{aligned} \tag{3.14}$$

Therefore, the expected degree for the entire network is twice the sum of the probabilities of existence of all the edges, divided by the number of nodes. This means that picking nodes at random from the network, we expect their degree to be equal to $\mathbb{E}[c]$ (on average) in eq. (3.14).

3.2.6 Clustering Coefficient

The clustering coefficient is defined as the fraction of path of length two that are closed. This coefficient quantifies the transitivity of the network. With transitivity we mean that if node i is connected to node j and node j is connected to node k, than also i is connected to k [262]. This property has fundamental implications on important network characteristics, such as the "small-worldness" [263]. The local clustering coefficient is calculated for each node in the network, but other definitions exist for a global measure of the transitivity. The most common way of defining the local clustering coefficient is the following

$$C = \frac{\#\text{ triangles } \times 3}{\#\text{ connected triples}} \tag{3.15}$$

where a "connected triple" is the configuration in which three nodes ijk are connected by the edges (i, j) and (i, k), whereas the edge (j, k) may be present or not. Since each triangle is counted three times when the triples ijk, jki, kij are evaluated, the number of connected triples is divided by 3.

In the case of a fuzzy network, as for the degree, the clustering coefficient of a node can take all the possible values with a certain probability. Therefore, also this feature must be redefined as a random variable. The probability of having a certain number of closed triangles in a triple depends on the probabilities of the corresponding edges and on the configurations of the edges in which that number of triangles occurs. For example, in the fuzzy network of fig. 3.1, each node may be tied to a triangle in 6 different ways. Precisely, in a network of N nodes, each node may be tied to t closed triangles in $\binom{N-1}{t}$ different configurations. Therefore, the probability of each configuration c can be computed as

$$q_c = P\left(\bigcap_{i,j\in S^c} e_{ij}\right) \cdot \left[1 - P\left(\bigcap_{i,j\in \bar{S}^c} e_{ij}\right)\right] \tag{3.16}$$

where S^c is the set of all the pairs of nodes (defining an edge) which define the configuration c, and $\bar{S}^c$ is its complementary. Since we are assuming that all the edges are independent Bernoulli random variables (eq. (3.7)), the intersection in section 3.2.6 can be taken out of the parentheses and replaced with the summation. The configurations can be considered as mutually disjoint and collectively exhaustive events, so that $\sum_c q_c = 1$. Consequently, the set of all configurations is regarded as the sample space of the network reconstruction experiment. In conclusion, the clustering coefficient probability distribution for the node i is given by

$$P_i^{cc}(C = \widetilde{C}) = \bigcup_{c\in\Gamma_i^{\tilde{C}}} q_c = \sum_{c\in\Gamma_i^{\tilde{c}}} q_c \tag{3.17}$$

where $\Gamma_i^{\tilde{C}}$ is the set of the configurations in which the node i has clustering coefficient equal to $\tilde{C}$. A representative example of the distribution is shown in fig. 3.1*b*, which shows the clustering coefficient distribution for each node of the depicted toy network.

3.2.7 Connected Components

Another fundamental feature of a complex network is the existence of a global connected component. A global connected component exists if there is at least one path from any node to any other node. Again, this feature is subjected to the stochasticity of the edges. The objective is to find the probability that all the nodes of the network belong to a unique connected component of k edges. To find the probability this we need to label the configurations which make the network completely connected with the related probability. Therefore, we can employ again the section 3.2.6 also to address the problem of the connectivity. In particular, we are asking for the probability that a network of N nodes, is completely connected by k edges. This probability reads

$$P_k^{cn} = \bigcup_{c \in \Gamma_{\mathbf{k}}^{cn}} q_c = \sum_{c \in \Gamma_{\mathbf{k}}^{cn}} q_c \tag{3.18}$$

where q_c comes from section 3.2.6 and $\Gamma_{\mathbf{k}}^{cn}$ is the set of the configurations in which exactly k edges make the network connected. These configurations can be efficiently found with a Breadth-First algorithm. The union sign on the left can be replaced by the summation because, as mentioned before, the set of the configurations is the sample space of the experiment, so that all the configurations are disjointed. An example is shown in fig. 3.1*c*, which illustrates the probability that the five nodes of the toy network belong to a single totally connected component with k existing edges (x-axis). Specifically, in the red window there are not enough edges to connect the network; in general, to do so are necessary at least $N-1$ edges. In the green window the probability increases and reaches the maximum. Finally, increasing further the number of edges required to connect the network, the probability decreases (orange window). It seems counter-intuitive that the probability of having a connected network decreases increasing the number of edges; the reason is that above a certain number of edges, the entire configuration becomes less likely, since the probability q_c of existence of all the k edges (at once) is smaller.

3.3 Numerical experiments and results

This section reports the results of the fuzzy network analysis for a set of synthetic and real-world networks. In both cases we considered undirected and unweighted

networks. In particular, we used 15 different synthetic network structures: 5 Erdős-Rényi, 5 Barabasi-Albert, 5 Watts-Strogatz, with 256 nodes each. The parameters of the generative models were fixed so that all the networks have expected degree approximately equal to 12. The three real-world networks considered are the collaboration network between Jazz musician [264], the food web of Little Rock Lake [265] and the brain network of the *Rhesus macaque* [266]. For each of the resulting 18 structures we generated 5 different dynamical realizations of two dynamical models: a linear (auto-regressive moving-average) ARMA(5,3)

$$x_t = \sum_{i=1}^{5} \alpha_i x_{t-i} + \sum_{i=1}^{3} \beta_i \varepsilon_{t-i} + \varepsilon_t + \gamma \tag{3.19}$$

where $\alpha_1, \ldots, \alpha_5$ and $\beta_1, \ldots, \beta_3$ are the model parameters for the auto-regressive and moving-average parts respectively, ε_t is a white noise random variable and γ is a constant; and a non-linear logistic model [267]

$$x_{n+1} = r x_n (1 - x_n) \tag{3.20}$$

where the parameter r is chosen randomly for each model realization in the interval $[3.57, 3.82]$ to assure a chaotic regime. Each time series spans a time horizon of 1024 time-steps. All the models take into account the connectivity of the underlying network using linear coupling terms, which are chosen to be small enough to guarantee that the resulting time series, especially in the case of the ARMA model, remain stationary. The result is a total of 180 numerical experiments.

The data used for the subsequent analysis are the time courses of the state variable of the nodes. Starting from this information we computed the connectivity matrix for all the networks using all the methods mentioned in Sec. section 3.2, obtaining a value of connectivity c_{ij} for each pair of nodes (i, j). The statistical significance of the connectivity was assessed by computing the corresponding *p-values* obtained by means of surrogate data analysis. Specifically, an adequate null hypothesis H^0_{ij} is the lack of relationships between the nodes i and j, which can be easily achieved by reshuffling the observed time course at each site [268]. The reshuffled time series possess the same mean, variance, and histogram distribution as the original signal, but any temporal correlation is destroyed, making this null model adequate to test for coherence or causal relations between nodes' dynamics. Nevertheless, other types of surrogates techniques can be used depending on the null hypothesis one would like to test. For instance, if one is interested in testing against the null hypothesis that the time series are correlated like in a random linear process, one should opt for Iterated Amplitude-Adjusted Fourier Transform-based surrogates, which preserve the linear features of the time series even in the frequency domain, while washing out higher-order dependencies.

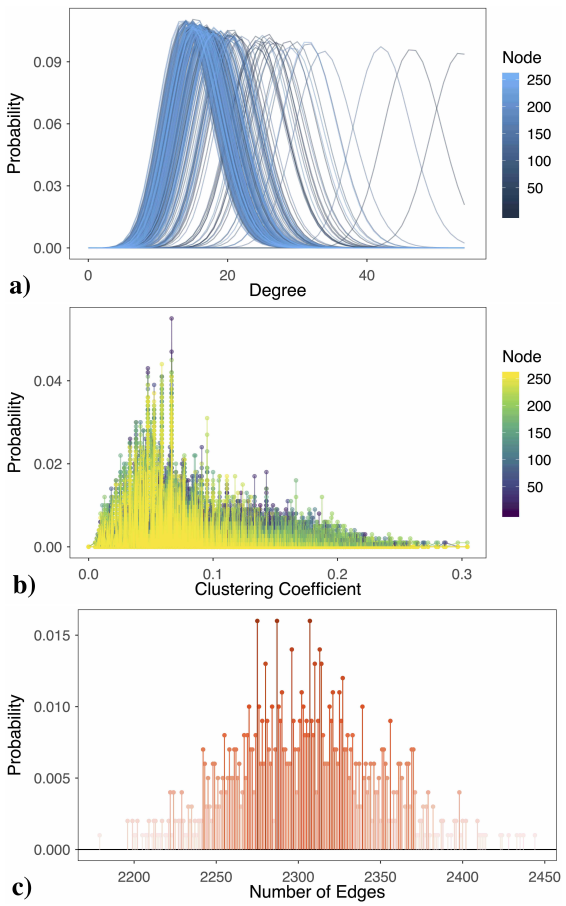


Figure 3.3 – A single realization of the ARMA dynamics on a single realization of the Barabasi-Albert network. Probability mass functions for the node degree (a), for the local clustering coefficient of the nodes (b) and for the connected component (c) as obtained from the fuzzy network analysis introduced in this study.

There are many additional types of surrogates that one can use to test against other null models, but the optimal choice is beyond the scope of the present paper. Changing null hypothesis could change the probability distributions of nodes' descriptors accordingly, but this neither negatively affects the goodness of the proposed method nor it can be easily related to the sensitivity of the method. The alternative hypothesis H^1_{ij} is that such relationships exist. If c^0_{ij} is the value of connectivity expected by chance for the edge e_{ij}, H^0_{ij} corresponds to $c^{obs}_{ij} = c^0_{ij}$, while H^1_{ij} is $c^{obs}_{ij} \neq c^{obs}_{ij}$. Given the empirical distribution of c^0_{ij} it was possible to obtain the *p-value* p_{ij} corresponding to the value of connectivity c_{ij} from the original time series. Subsequently, the resulting *p-values* were used to obtain the minimum posterior probabilities through section 3.2.2 and section 3.2.2. As mentioned in Sec. section 3.2, the only free parameter of the process is the prior probability for the null hypothesis $P(H^0_{ij})$. In this experiment, we set it equal to $1 - D$ for all the edges, where D is the average density of the networks considered. With this choice we are allowed to write the prior probability as $P(H^0)$, without the subscript ij. This is clearly not the optimal choice for the prior probability, since it is not true that it is equivalent for all the edges. However, this puts us in a scenario where only a global prior information about the network structure is available. The computation of the minimum posterior probability for all the p_{ij} returns the adjacency matrix $\mathbf{P}$ of the fuzzy network, which is used to compute all the network descriptors as described in the previous section. It is to be noticed that, in general, the prior $P(H^0_{ij})$ value determines a shift in the probability π_{ij} of $\mathbf{P}$ in the interval $[0, 1]$; the same

happens in the particular case in which $P(H^0_{ij}) = P(H^0)$. Unlike the traditional methods for network reconstruction, which use heuristics to determine a threshold on c_{ij} or on p_{ij}, our process maintains the uncertainty on the parameters until the computation of the network descriptors.

Figure fig. 3.3 shows an example for the realization 1 of the Barabasi-Albert network, using the Pearson correlation coefficient with ARMA dynamics. The three figures are analogous to those in fig. 3.1 for the toy network. The fig. 3.3a) shows the probability mass functions for the degree of each node, which follow the Poisson-Binomial distribution in eq. (3.8). We used a free R routine for the numeric approximation of the analytical distribution provided by [269]. It is to be noticed that the vast majority of the distributions are grouped around the real average degree of the network, that is 12.8. Our theoretical median prediction in this case is 15.03 with $[11, 21]$ 68% confidence interval.

Similarly, fig. 3.3b) shows the probability mass functions for the local clustering coefficient of the nodes. The distributions are very irregular and do not follow any known probability function. For comparison, the average clustering coefficient in the real network is 0.107 whereas our theoretical median prediction is 0.0603 with $[0.0521, 0.1116]$ 68% confidence interval.

Finally, in fig. 3.3c) is reported the probability mass function for the connectivity, which represents the probability that all the nodes of the network belong to the same connected component of k edges. The distribution overestimates the number of edges needed to have a unique connected component, since the ground-truth network, which is actually connected in one component, consists of 1515 edges. A possible explanation for this overestimation is that the π_{ij} might be too high due to the choice of the prior $P(H^0)$ in section 3.2.2, whose effect is to shift the values on the vertical axis.

A useful synthesis of these network descriptors (besides the expected value) is the maximum posterior probability (MPP) estimate, which is simply the mode of the computed distributions. Furthermore, other statistics about the dispersion can be computed to assess the uncertainty on the values at node level. This is made possible by the fuzzy approach which does not require any threshold – neither on the connectivity matrix nor on the *p-values* – allowing the uncertainty to be considered as part of the network analysis, rather than an obstacle to overcome.

Figures fig. 3.4 and fig. 3.5 show an overview on the results of the whole set of numerical experiments. The figures include the comparisons between the fuzzy procedure and another well-known method, which consists of thresholding the connectivity matrix $\mathbf{C}$ (section 3.2.1) to derive the binary adjacency matrix. Several criteria exist to fix a value for the threshold, (see e.g. [270, 271]). Instead, for the sake of comparison, we reconstructed the networks with several threshold levels, according to the range of the connectivity provided by each tool. In

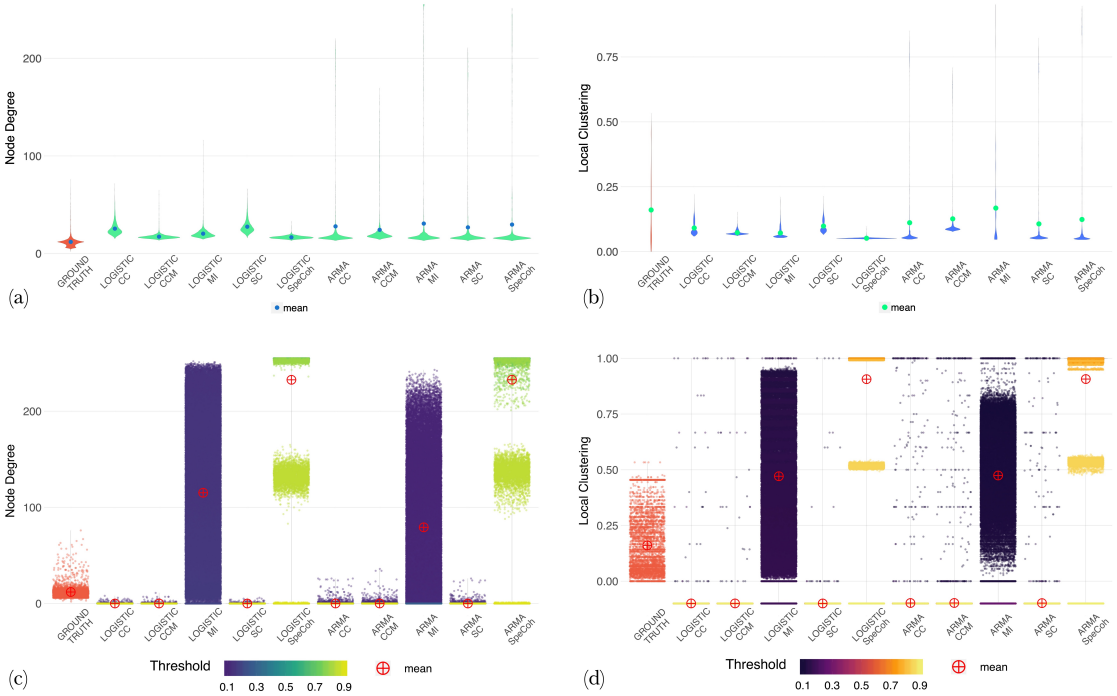


Figure 3.4 – **Summary distributions** for the degree (a) and the local clustering coefficient (b) for all the synthetic network, grouped by dynamics and connectivity measures. The figures (c) and (d) show the distributions of the degree and the local clustering coefficient by varying the threshold level, which is encoded by the color, and with respect to different dynamics (ARMA or LOGISTIC) and discriminating statistics for network reconstruction (CC, CCM, MI, SC, SpeCoh; see the text for details). Each point corresponds to a single node (a random horizontal jitter is added): note that the same node appears multiple times with different colors, for each value of the threshold. In every plot it is also indicated the mean of the distributions.

particular we used equally spaced values in the interval $[-1; 1]$ for the statistical tools (CC, SC, SpeCoh) and for the Convergent Cross Mapping (CCM), whereas equally spaced quantiles were supplied for the Mutual Information (MI). The plots in Fig. fig. 3.4a–d are obtained from the aggregating the 150 synthetic network experiments, while those in Fig. fig. 3.5a–d are aggregated over the 30 real network experiments. The results are grouped according to the dynamic and the discriminating statistics used to assess the connectivity. Each point in Fig. fig. 3.4c–d and fig. 3.5c-d represents the value of degree and clustering for a single node given a threshold value for the connectivity matrix.

Overall, analysis shows that the majority of the results obtained from the fuzzy analysis are consistent with the expectations, whereas the thresholding approach, regardless of the statistical method, tends to underestimate or overestimate the true values for varying thresholds. These results provide a strong indication that results from threshold models not only strongly depend on the value of the thresh-

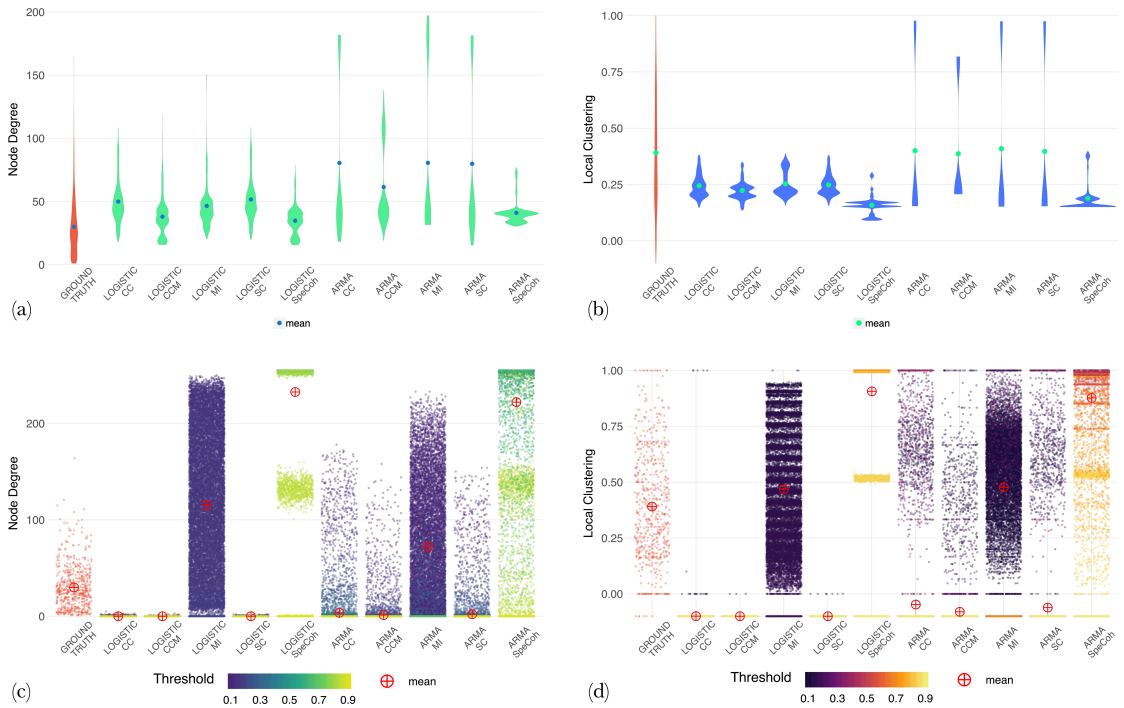


Figure 3.5 – **Summary distributions** for the degree (a) and the local clustering coefficient (b) for the real-world network, grouped by dynamics and connectivity measures. The figures (c) and (d) show the distributions of the degree and the local clustering coefficient by varying the threshold level, which is encoded by the color, and with respect to different dynamics (ARMA or LOGISTIC) and discriminating statistics for network reconstruction (CC, CCM, MI, SC, SpeCoh; see the text for details). Each point corresponds to a single node (a random horizontal jitter is added): note that the same node appears multiple times with different colors, for each value of the threshold. In every plot it is also indicated the mean of the distributions.

old, but also that there can be no thresholds for which, on average, reliable measure of network indicators as simple as degree centrality and local clustering coefficient can be obtained.

Let us discuss in greater detail the results concerning the fuzzy network analysis. All the discriminating statistics yield meaningful results in terms of expected value although, in the considered cases our method slightly overestimates the average degree. The clustering coefficients instead, show better average values for the linear dynamics, despite a general slight underestimation. Both types of deviation from expected values of the features are due to the shift effect of $P(H^0)$. In this specific setting, the lower values of the two descriptors are not well captured, because every edge has a positive – albeit very low – probability of existence which keeps the degree and the clustering away from zero.

As expected, the fuzzy network analysis applied to ARMA dynamics returns

the less valuable outcomes, arguably because of the very low coupling that we imposed to the edges. This result is still consistent with our expectations: a low value of the coupling results in statistical correlations more difficult to detect even in the case of linear dynamics. In this case, because of this additional source of uncertainty due to such a limitation, the inferred values span a broader interval for both degree and clustering coefficient: nevertheless, most of the mass meets the ground-truth distribution, suggesting that the fuzzy network analysis is able to robustly cope with the increased level of uncertainty. The best tool to derive the connectivity proved to be the CCM, which allows the non-linearity to be taken into account adequately. It is worth noting that the performance of statistical methods and the overall results may be improved by adjusting for the spurious relationships such as the partial correlations, but a direct implementation of this task is beyond the scope of the present work.

Remarkably, the fuzzy descriptors outperform the traditional thresholding reconstruction methods in all the cases. Despite the numerous threshold levels in place, the real values of the network features are rarely detected by the latter reconstruction technique. In some cases, the results reflect the ground-truth, but only for specific values of the threshold which remain basically arbitrary. The results for the real-world networks are qualitatively analogous (fig. 3.5), whereas the performances for the ARMA models have improved for both the network descriptors. The real-world networks span a broader range of values for both the node degree and the clustering coefficient. This is clearly reflected in both the methods presented. Even in this case, the distribution of the two network descriptors are skewed towards the lower values; this feature is mostly captured by the fuzzy model, while just specific values of the threshold accomplish the ground-truth.

3.4 Future directions

Here we provide a brief discussion about possible future research directions. In fact, our work opens the way to the definition of other descriptors – e.g. centrality measures – of complex systems in the wake of the fuzzy descriptors. Also, the fuzzy perspective might be extended to the dynamical features of a complex network by studying, for instance, the properties of the fuzzy counterpart of the Laplacian.

A first example of another topological descriptor eligible to be redefined in fuzzy terms is the rich-club coefficient (see e.g. [272]). Using the information coming from the degree distribution for the individual nodes, the rich-club coefficient can be redefined as the probability $P^{rich}(e, n, k)$ to observe e edges connecting n nodes of degree greater than k. Having the Poisson-Binomial probability distribution for the degree (eq. (3.8)) and the probabilities of the possible configurations (section 3.2.6), all the ingredients are there to obtain P^{rich}. Given a fuzzy net-

work of N nodes and the associated fuzzy adjacency matrix with elements π_{ij} the probability that the node i has degree greater than k is

$$P^0 = P(d_i > k) = 1 - CDF(\mathfrak{P}[k, \pi_{i\cdot}]) \tag{3.21}$$

where $j \in 1, ..., N$ and $j \neq i$ and $\mathfrak{P}[k, \pi_{ij}]$ is the Poisson-Binomial distribution of parameters $\overrightarrow{\pi}_{i\cdot}$ (the row i of the fuzzy adjacency matrix). Let's select n nodes with $n \in 2, 3, ..., N$ from the possible $\sum_{n=1}^{\infty} \binom{N}{n}$ configurations. The configurations are indexed as c with $c \in \left[1, \binom{N}{n}\right]$. The nodes selected in the particular configuration c form a set S_c with cardinality $|S_c| = n$. The probability that *all* the n nodes $i \in S_c$ of the configuration c have degree greater than k is

$$P_c^{>k} = P(d_{S_c} > k) = \prod_{i \in Sc} P_c^0 = \prod_{i \in Sc} P(d_i > k) \tag{3.22}$$

under the hypothesis of independence. Finally, the probability of existence of e edges between the nodes of degree greater than k is given by

$$P_c^{rich} = P(E_{>k} = e) = P_c^{>k} \cdot \mathfrak{P}[e, \overrightarrow{\pi}_c] \tag{3.23}$$

where $\overrightarrow{\pi}_c$ is the vector of probabilities of existence of the edges in the configuration c. This probability makes use again of the Poisson-Binomial distribution as it is capable to model the presence of the edges in the configurations. To obtain the probability distribution $P^{rich}(e, n, k)$ for the rich-club coefficient of the network, regardless of the configuration, the inclusion-exclusion criterion must be applied to the above equation.

Leveraging on the descriptors showed so far, it is possible to compute the probability of observing a random walk occupying a particular node. The probability is usually given by $\frac{k}{2m}$: in the case of a fuzzy network it is replaced by the ratio distribution of a Poisson-Binomial and a Gaussian distribution (see eq. (3.8) and section 3.2.5).

The information about the degree distribution can also be exploited to study the fuzzy counterpart of assortativity of the network, for example starting by defining the excess degree distribution for the single node and consequently for the entire network.

We expect that further analysis in this direction, left for future studies, will lead to interesting results from both theoretical and applied perspectives.

3.5 Conclusions

In this chapter, we have presented a novel framework for network analysis under uncertainty about the underlying connectivity, which overcomes some of the

issues typical of network reconstruction procedures. The proposed framework is fully general, and it is not restricted to the study of mobility systems, but may have valuable application also in network neuroscience, environmental science and whenever a robust knowledge of the topological features of an uncertain network is deemed important. Our approach can be used to infer the structural features of a complex system when its topological connectivity is specified by edge probabilities. Also, we proposed a simple method, mathematically grounded, for the computation of these probabilities from a set of p-values, usually obtained from one's preferred analytical technique.

The leading idea is to define a new standpoint, from which the uncertainty about the structural features of a complex system can be detected and quantified, without the explicit reconstruction of the underlying network structure. The method does not require any assumption on the topology of the network under study, nor on the underlying generative process. Conversely, our approach enables one to include prior knowledge about the existence of the single edges in a rigorous manner, without fixing any arbitrary threshold nor the level of statistical significance. Starting from the *p-values* associated with measures of correlations or statistical causality, we obtained a Bayesian definition of the probability of existence of the single edge. The probabilities are rearranged in an adjacency matrix, which represents a "fuzzy" model used to elicit relevant information on the network structure. All the information have a stochastic nature which allows uncertainty to be assessed. Consequently, we proposed new definitions of some important network descriptors such as the node degree and clustering coefficient, considering them as random variables. Finally, we compared the results with a very well-known method for network reconstruction, showing the strength and weakness of our procedure. From a computational perspective, for small networks the probabilities of the configurations can be directly computed with section 3.2.6, which gives all the information to derive the probabilities for the clustering coefficient and the connectivity. In case of large networks the computational effort might be prohibitive: however an adequate sampling procedure from the fuzzy adjacency matrix in section 3.2.3 can be performed to compute the distributions and the statistics of interest. The method may be enhanced by applying different tools which take into account the partial relations in deriving the connectivity matrix (e.g. partial correlations, PCMCI algorithm). In addition, past studies can be readily integrated with our approach given the *p-values* obtained therein.

Appendix

From *p-value* to probability

In general, the *p-values* can be formally considered as random variables (see e.g. [273]), which under the null hypothesis are distributed uniformly in $[0, 1]$. This is a direct consequence of the Probability Integral Transform: given the test statistic T of interest, and its realization t, the *p-value* is by definition (using the same notation as in the manuscript)

$$\begin{aligned} p_{ij} = P\left(T \geq t \mid H_{ij}^0\right) &= 1 - P\left(T < t \mid H_{ij}^0\right) \\ &= 1 - F_T^0(t) \end{aligned}$$

All the subsequent formula are intended under the null hypothesis, so we drop the 0 at the apex. Let's now define the random variable $U = F_T(t)$ (see also fig. 3.6). It follows that

$$\begin{aligned} F_T(t) = P\left(U \leq u\right) &= P\left(F_T(t_u) \leq u\right) = P\left(T \leq t_u\right) \\ &= P\left(T \leq F_T^{-1}(u)\right) = F_T\left(F_T^{-1}\left(u\right)\right) \\ &= u \end{aligned}$$

which is equivalent to the definition of a Uniform distribution for the variable U. Since $p_{ij} = 1 - F_T(t) = 1 - U$ we need to prove that also $F_P(p_{ij}) \sim \text{Unif}(0, 1)$:

$$\begin{aligned} F_P\left(p_{ij}\right) = P\left(P \leq p\right) &= P\left(1 - U \leq p_{ij}\right) \\ &= P\left(U \geq 1 - p_{ij}\right) \\ &= 1 - P\left(U \leq 1 - p\right) \\ &= 1 - 1 - p \\ &= p \quad \square \end{aligned}$$

It is to be noticed that, under the alternative hypothesis, the *p-values* are not uniformly distributed, but are typically skewed. Therefore, as we explained in the manuscript, the distribution of the *p-value* p_{ij} (for the edge e_{ij}) is modeled as $\text{Beta}(\xi, 1)$ following the procedure of [258]:

$$p_{ij} \sim f(p_{ij}|\xi) = \xi p_{ij}^{\xi - 1}. \tag{3.24}$$

The standard Uniform distribution is a particular case of section 3.5, where the parameter $\xi = 1$

$$\text{P}\left(p_{ij}|H_{ij}^0\right) = f(p_{ij}|\xi = 1) = 1$$

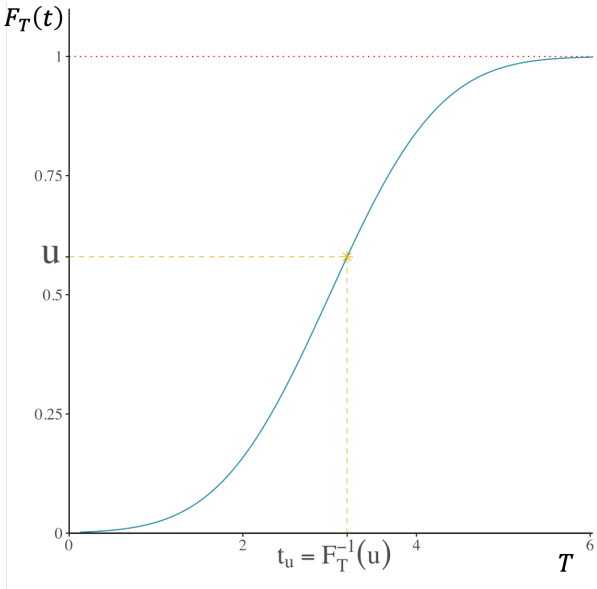


Figure 3.6 – Probability Integral Transform example

The section 3.2.2, in the main text, reports the definition of the Bayes Factor, from which we want to prove that a lower bound for the odds of H_{ij}^0 on H_{ij}^1 is represented by section 3.2.2, synthetically reported here:

$$B_{ij} = \inf_{\xi} B_g(p_{ij}) = -ep_{ij} \log p_{ij} \quad \text{for} \quad p_{ij} < e^{-1}$$

and equal to 1 otherwise.

Since the numerator of section 3.2.2 is a constant, the lower bound for B_{ij} corresponds to the upper bound of the denominator:

$$B_{ij} = \inf_{\xi} B_g(p_{ij}) = \frac{1}{\sup_{\xi} \int_0^1 f(p_{ij}|\xi) g(\xi) d\xi} \tag{3.25}$$

From the First Mean Value Theorem we know that in general, if $f : [a, b] \to R$ is continuous and g is integrable and does not change sign on $[a, b]$, then there exists some c in (a, b) such that

$$\int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx$$

In our case we have that

$$\begin{aligned}\int_0^1 f\left(p_{ij} \mid \xi\right) g(\xi) d\xi &= \int_0^1 \xi p_{ij}^{\xi-1} \cdot g(\xi) d\xi \\ &= \bar{\xi} p_{ij}^{\bar{\xi}-1} \int_0^1 g(\xi) d\xi \\ &= \bar{\xi} p_{ij}^{\bar{\xi}-1}\end{aligned}$$

Thus we can rewrite the section 3.5 as

$$B_{ij} = \inf_{\xi} B_g\left(p_{ij}\right) = \frac{1}{\sup_{\bar{\xi}} \bar{\xi} p_{ij}^{\bar{\xi}-1}} \tag{3.26}$$

We now define $h\left(\bar{\xi}\right) = \bar{\xi} p_{ij}^{\bar{\xi}-1}$, so that

$$\begin{aligned}h'(\xi) = p^{\bar{\xi}-1} + \bar{\xi}\left(p^{\bar{\xi}-1} \ln p\right) = 0 \\ p^{\bar{\xi}-1} = -\bar{\xi} \ln(p) \cdot p^{\bar{\xi}-1}\end{aligned}$$

which is true only if $\bar{\xi} = -\frac{1}{\ln p}$. Substituting this result in section 3.5 we obtain

$$\begin{aligned}B_{ij} = \frac{1}{-\frac{1}{\ln p} \cdot p^{-\left(\frac{1}{\ln p}+1\right)}} &= -\ln p \cdot p^{\frac{1}{\ln p}} \cdot p \\ &= -p \ln p \cdot p^{\log_p e} \\ &= -ep \ln p \quad \square\end{aligned}$$

Note that, since for $p > e^{-1}$ the function B_{ij} is decreasing, we impose that for *p-values* larger than $e^{-1} \approx 0.368 \rightarrow B_{ij} = 1$.

Given the definition of the Bayes Factor as

$$B_g\left(p_{ij}\right) = \frac{\mathrm{P}\left(p_{ij} \mid H_{ij}^0\right)}{\mathrm{P}\left(p_{ij} \mid H_{ij}^1\right)} = \frac{\mathrm{P}\left(H_{ij}^0 \mid p_{ij}\right) \cdot \mathrm{P}\left(H_{ij}^1\right)}{\mathrm{P}\left(H_{ij}^1 \mid p_{ij}\right) \cdot \mathrm{P}\left(H_{ij}^0\right)}$$

it follows that

$$\begin{aligned}\mathrm{P}\left(H_{ij}^0 \mid p_{ij}\right) &= B_{ij} \cdot \frac{\mathrm{P}\left(H_{ij}^0\right) \mathrm{P}\left(H_{ij}^1 \mid p_{ij}\right)}{1-\mathrm{P}\left(H_{ij}^0\right)} \\ &= B_{ij} \cdot \frac{\mathrm{P}\left(H_{ij}^0\right)\left(1-\mathrm{P}\left(H_{ij}^0 \mid p_{ij}\right)\right)}{1-\mathrm{P}\left(H_{ij}^0\right)}\end{aligned}$$

which finally gives

$$1-\pi_{ij} = \left(1+\left(\frac{B_{ij} \cdot P(H_{ij}^0)}{1-P(H_{ij}^0)}\right)^{-1}\right)^{-1}$$

which is the section 3.2.2 of the main text.

Lyapunov condition for the Poisson - Binomial distribution

In section section 3.2.4 we defined the fuzzy degree of a single node, which is a random variable following the Poisson-Binomial distribution. In order to obtain the expected degree of a network we relied on the fact that the Poisson-Binomial converges to the Normal distribution if the Lyapunov condition were satisfied. Here we prove that the condition is satisfied under very broad conditions.

Let $d_i \sim \text{Bernoulli}\,(p_i)$, with $d_1, d_2, \ldots$ independent but not identically distributed random variables, represent the degree of node i as stated in eq. (3.7). Let also $X_i = d_i - \mu_{d_i} = d_i - p_i$. Defining $s_n^2 = \sum_{i=1}^n \sigma_{d_i}^2$ we can rewrite the Lyapunov condition as

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E}\left[|X_i|^{2+\delta}\right] = 0 \Longrightarrow \frac{1}{s_n} \sum_{i=1}^n X_i \xrightarrow{d} N(0,1).$$

We prove that the Poisson-Binomial probability distribution satisfies this condition, by finding an upper bound converging to zero to the above sum.

To do so we observe that

$$1 \geq p_i\,(1-p_i) = \sigma_{d_i}^2 = \mathbb{E}\left[X_i^2\right] \geq \mathbb{E}\left[|X_i|^{2+\delta}\right]$$

for any $\delta > 0$. Therefore,

$$\begin{aligned}
&\frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E}\left[|X_i|^{2+\delta}\right] \leq \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E}\left[X_i^2\right] \\
&= \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \sigma_{d_i}^2 = \frac{1}{s_n^{\delta}}
\end{aligned}$$

Consequently, since $s_n \to \infty$ (exept for degenerate cases where $p_i = 0$ or $p_i = 1$ for all i), the Lyapunov condition is satisfied and similarly it is the "normalized" Poisson-Binomial random variable follows $\sum_{i=1}^n X_i/s_n \xrightarrow{d} N(0,1)$.

Chapter 4

The effects: COVID-19 lockdown unravels the complex interplay between environmental conditions and human activity

In brief

During COVID-19 epidemic, draconian countermeasures forbidding human mobility and non-essential human activities have been adopted in several countries worldwide, providing an unprecedented setup for quantifying their *effects* on the environment. Here, we unravel the causal relationships between 16 variables – including different flavors of human mobility flows – considered as the components of a complex socio-environmental system, and apply information theory, network science and Bayesian inference to map the backbone of the complex interplay between them. We introduce a novel information-theoretic method based on statistical divergence to identify abrupt changes in the system dynamics, caused by a sudden intervention. We find that despite a measurable decrease in NO_2 concentration, locking down a region may be an insufficient remedy to reduce emissions. Our results provide a functional characterization of socio-environmental interdependent systems.
This chapter is published in *Complexity* [18].

4.1 Introduction

Εν το παν

Chrysopoeia of Kleopatra

Complex systems consist of interconnected units which are characterized by nonlinear dynamics at the microscopic scale that lead, at larger scales, to emergent collective phenomena such as human mobility and environmental changes. Often, those units are complex sub-systems interacting with each other, exhibiting a multilayer [274] or interdependent organization [275]. Usually, the study of human mobility analyses the role of drivers, such as environmental conditions, on mobility patterns; here conversely, we bring to attention the *effects* of mobility on the surrounding environment, following the complexity science perspective. Complex networks, in particular, provide an abstract representation for the backbone of such systems. In fact, it has been shown that shocks in one node or in one sub-system can quickly propagate to the rest of the network, driving the overall system to a catastrophic collapse [276, 277, 278]. However, understanding the resilience of a complex system can be even more challenging when its backbone cannot be directly observed and must be inferred from indirect observations, such as the time course of a set of physical observables. In fact, the collective behaviour of the system is driven by the interplay between the dynamics at each component and the hidden interactions among them. In statistical physics the reconstruction of the system's interactions from measuring time series related to its components is usually referred to as inverse problem [279]. In the last decades a great effort has been devoted to the solution of this problem [280, 281, 229, 239, 241], that has relevant implications in various disciplines, from neuroscience [282], to ecology [283], including finance [212], biology [284] and climatology [255]. In chapter 3 we presented a new general framework to analyse the *structure* of a complex system when the underlying network is either hidden or uncertain and the focus was on the topological descriptors of the system. In this chapter, instead we are dealing with the *effects* of a change in human activities, including mobility, on the surrounding environment. The structural relationship between environmental, climate and human-dependent variables cannot be directly measured, however, using the methods presented in section 2.2 we map the functional backbone of these interdependent sub-systems to understand their response, and consequently their resilience, to exogenous and endogenous perturbations.

To this aim, we focus on a specific case study, namely the reduction of pollutants in the Northern Italy, observed between March and July 2020 in response to draconian interventions due to COVID-19 pandemic. In fact, during the first months of 2020, Italy has been one of the country mostly affected by

COVID-19 [285, 286, 287, 288, 289, 290, 291, 292]. To prevent the spread of the SARS-CoV-2 virus, Italy adopted significant non-pharmaceutical interventions [293, 294, 295, 296], including locking down the entire country from 9th of March to 4th of May. The forced closure of school, public facilities and places of employment drastically reduced mobility, with the most relevant effects in the Lombardia region, in the North of Italy, which is also the Italian region most plagued by the virus. On the one hand, this unprecedented situation triggered a cascade of public health, social, behavioral and economic challenges that will require years to recovery. On the other hand, from a scientific perspective, one can investigate the effects of such dramatic regional and sub-regional interventions on the environment. In practice, the spread of COVID-19 allows one to better understand the causal and functional relations between the components of the socio-environmental system, also shedding light on the potential effects that large-scale interventions, such as mobility restrictions, may have on the system's collective behaviour. This insight would also be a valuable resource to design informed policies that can be adopted to mitigate the climate crisis. To this aim, we consider this situation as a global experiment, giving us a unique opportunity to investigate the complex interactions between human activities and environment with a particular focus on human mobility and air quality, using the North of Italy as a case study where the availability of heterogeneous data sources allows one to perform a more integrative analysis. The interest in such a relationship has exploded worldwide during the lockdown period and many recent studies focus on the reduction of atmospheric pollutants due to mobility restrictions in different countries [297, 298, 299, 300, 301]. However, some works highlight that reductions in human mobility and in industrial activity would not be sufficient to reduce air pollution, especially when meteorology is unfavorable [302]. In fact, it must be noticed that air quality does not only depend on emissions (and consequently concentrations) of pollutants but it is strongly influenced by meteorology, which plays a crucial role in the Po Valley, where Lombardia region is located. Moreover, the effect of the topography, with the Alps to the north, contributes in making this region one of the most polluted areas around Europe.

Here, our purpose is to unravel the complex interactions between environmental conditions, human mobility and energetic consumption, taking the Lombardia Region during the Italian lockdown as a case study. Mapping the network structure of a socio-environmental system is crucial to understanding its collective behaviour and to acknowledge the emergence of unexpected outcomes [303, 304]. To this aim we reconstruct the network of functional and causal relations between the observables, and subsequently identify the (causal) influence that an external intervention – as the lockdown – exerts on the system. In particular, we aim at evaluating the impact of relaxing non-essential human activities – i.e., those activ-

ities not directly related to the supply of goods and commodities – on tropospheric NO_2 concentrations during the lockdown. We employ representative data for environmental conditions – i.e air pollution and meteorological conditions data – and human activity – i.e. mobility and energy load data – over a survey period of four months – February, March, April and May – both in 2019 and in 2020, for a total of 16 variables (see fig. 4.1). By taking into account all these variables we reduce the possible disturbances due to latent confounders. Afterwards, we evaluated the differences in the dynamics of such variables between the period related to a situation with extremely reduced activity and baseline periods. Since the lockdown imposed social distancing and, consequently, drastically reduced human mobility, we referred to an indicator of air quality which is sensitive to mobility changes, i.e. the *nitrogen dioxide* (NO_2) concentrations, which proves to highly depend on emissions from transportation means and industrial activity [305]. Concerning the meteorological conditions, we considered the daily average of 6 variables – i.e surface pressure, surface net solar radiation, temperature, total precipitation, wind direction and wind speed. We used the data made publicly available by

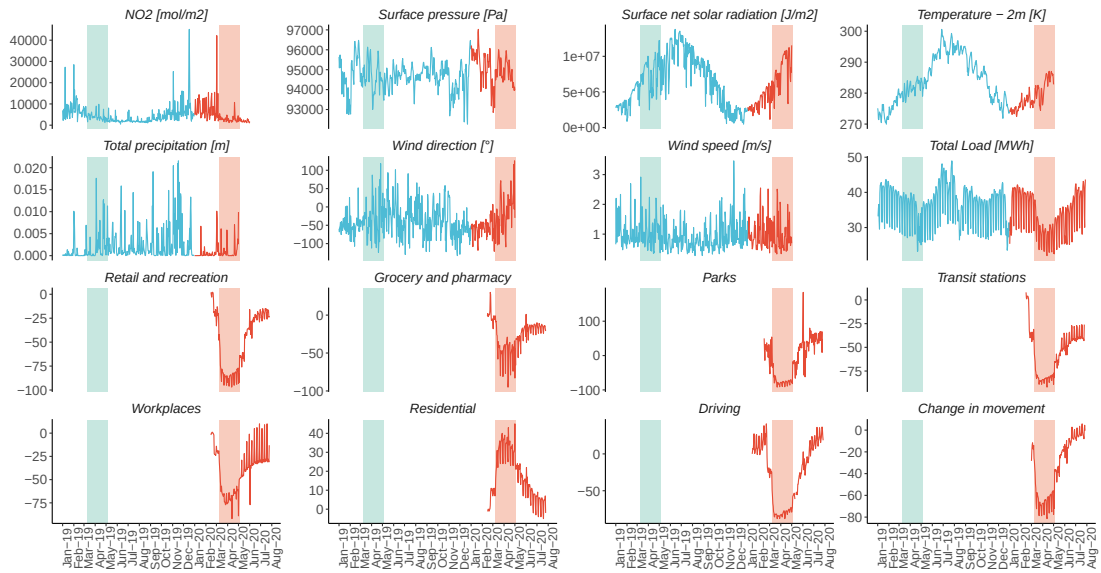


Figure 4.1 – Observables used as a proxy for measuring variations in environmental conditions and human activities during the Italian lockdown. Panels show the time series for each of the 16 variables considered in this study; blue and red curves corresponds to the times courses for year 2019 and 2020, respectively. The light-red band corresponds to the lockdown period in 2020, while the light-blue band corresponds to the same period in 2019. The public mobility data are available only for the year 2020.

Google [306], Apple [307] and Facebook [308] during the survey period to investigate human mobility changes. In Lombardia region, energy consumption is mainly attributable to industry, consequently we have considered the daily average electricity system's total load, as a proxy of industrial activity (see *Appendix 4.4* for details).

From a methodological point of view, we evaluated the significance and the concomitance of variations in human activities and in environmental conditions – focusing on NO_2 concentrations – in the survey period 2019-2020, using statistical and causal analysis. Through statistical tests and effect size measures [309, 310] we were able to assess the relevant variations in the time series of the considered variables. To investigate the complex nexus between the 16 variables we leveraged on the partial correlation coefficient (PCC) [311] and Granger causality (GC) [181]. These two measures provide a functional and causal topological maps of the complex system. Finally, we assess the causal impact of the lockdown intervention by taking inspiration from a relatively recent Bayesian technique, based on a state-space model, used to infer the causal impact of advertising campaign on the market sales [165]. By specifying which period in the data should be used for training the state-space model (pre-intervention period) and which period for computing a counterfactual prediction, this technique assesses the impact of the attributable intervention [165].

4.2 Identifying tipping points in empirical observations

To determine whether the data reflect the regime shift of the lockdown we implemented a shifting point detection technique based on an information-theoretic measure of similarity, called Jensen-Shannon Divergence. The Jensen-Shannon Divergence (JSD) is an information theoretic measure which quantifies the difference between two probability distributions. It has the main advantage of being symmetric and that its square root defines a metric: therefore, it can be used like a distance measure to quantify (dis)similarity. Given two discrete distributions, $P(x)$ and Q(x), over the same probability space – $\mathbb{R}$ in our case – one can define the Kullback-Leibler divergence (KLD) from Q to P as

$$D_{KL}(P||Q) = \sum_x P(x) \log_2 \frac{P(x)}{Q(x)}. \tag{4.1}$$

This measure is also known as relative entropy and, as by definition, it is not symmetric. It measure the amount of bits that one gains about using $Q(x)$ to model the distribution $P(x)$: in fact, when $Q = P$ then the divergence is zero bits.

This relative entropy has also the disadvantage of not being upper bounded: for this reason, other measures like the JSD are also widely used. By introducing the mixture distribution $\mu(x) = (P(x) + Q(x))/2$, the JSD can be defined in terms of the KLD as

$$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||\mu) + \frac{1}{2}D_{KL}(Q||\mu). \tag{4.2}$$

It can be shown that $0 \leq D_{JS}(P||Q) \leq 1$ bits. In our study, the role of P and Q is played by two sub-periods of observation: given a time series $s(t)$, it is split into $s_1(t)$ for $t \in [t_0, T]$ and $s_2(t)$ for $t \in [T+1, t_f]$, being t_0 and t_f the initial and final observation time, respectively. Here, T represents the instant of time at which the observation series switches between two regimes. The time instant T is determined by varying its value and computing the divergence D_{JS} until it becomes statistically significant.

The result of the process is the date of the tipping point in which the dynamics meets a regime shift, splitting the survey period in pre-lockdown and lockdown periods. This subdivision is used both for the statistical tests and for the Bayesian state-space model presented in the following. The date of the "information-theoretic" lockdown turns out to be the 14th of March (p-value $= 6.12 \cdot 10^{-5}$), 5 days after the institutional lockdown. This result is consistent with the physical behaviour of the NO_2 which has a typical lifetime of few days. Consequently, we grouped the time series in the 4 subsets used for testing:

- Group 1.1 (pre-lockdown) from 1 February 2019 to 14 March 2019
- Group 1.2 (lockdown) from 14 March 2019 to 5 May 2019
- Group 2.1 (pre-lockdown) from 1 February 2020 to 14 March 2020
- Group 2.2 (lockdown) from 14 March 2020 to 5 May 2020

It is to be noticed that in 2019 there was no lockdown, but for sake of simplicity in the following we will refer to the Groups 1.1 and 1.2 as pre- and post-lockdown periods *for 2019* since they corresponds to the same periods of the year of the actual intervention of 2020.

Figure 4.1 provides an overview of the time course of each observable used during the entire survey period. In fig. 4.2 is reported in detail the NO_2 variation, comparing the two years of reference, suggesting that the average NO_2 concentration during the lockdown period in 2020 is smaller than the one during the same period in 2019. In the following this observation will be corroborated by statistical analysis. It is worth noticing that, in general, at the end of the winter season the NO_2 concentration is expected to decrease due to the reduction of the domestic,

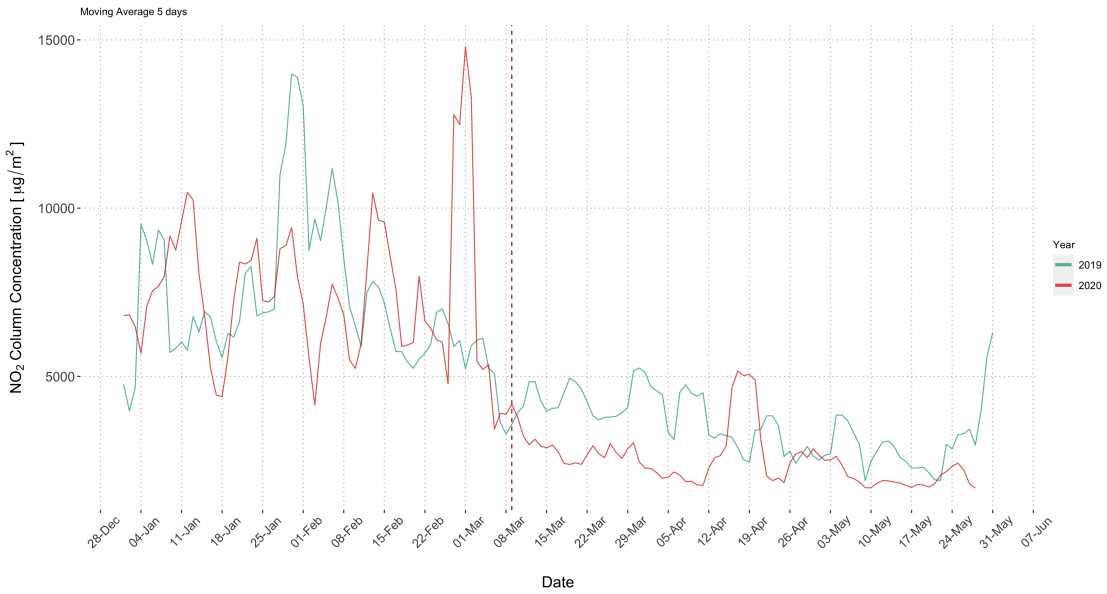


Figure 4.2 – Comparison of NO_2 concentration in 2019 and 2020. The figure shows the 5-days moving-average concentration of the NO_2 in the two years of reference. The dashed vertical line indicates the beginning of the lockdown period on 9 March 2020

civil and industrial heating concurrently with the rising temperatures. In 2020 the reduction appears to be more abrupt with the beginning of the lockdown and the concentration levels remain lower for the rest of the period. The variability of the concentration during the lockdown is also lower than the one in the same period in 2019 (Fligner test on the difference of variances [312] gives a p-value $= 3.514 \cdot 10^{-7}$) this may be due to the reduction in the fast-changing pressure variables, such as transportation. The results of t-tests [313] support the visual assessment of fig. 4.2, confirming that even though the NO_2 concentration in the pre-lockdown periods (in 2019 and in 2020) are statistically equivalent (the hypothesis of equal average concentration for the two periods cannot be rejected with a t-test p-value $= 0.53$), the reduction observed during the 2020 lockdown with respect to the same period in 2019 is statistically significant (the average concentration in the two periods are different, with a t-test p-value $= 7.39 \cdot 10^{-4}$). These results confirm that the NO_2 average concentration in 2020 were significantly lower than in 2019. Now we want to evaluate also the magnitude of the observed differences by computing their effect-size. Two commonly used measures of the effect size are the Cliff-δ and the C.L.E.S. (Common Language Effect Size) [314, 315] (see also the *Appendix* at the end of this chapter). For the pre-lockdown periods, these measures indicate a very low effect size ($\delta \simeq 0$ and $CLES \simeq 0.5$), whereas it is very high for the post-lockdown periods ($\delta \simeq 0.7$ and $CLES \simeq 0.806$). The CLES measure provides

an immediate interpretation of the results, in particular, a $CLES \simeq 0.806$ means that the probability that an observation from the 2020 lockdown period returns a lower value of NO_2 concentration w.r.t. the same period in 2019 is more than 80%. A complete summary of the statistical tests is reported in table 4.2 of the section 4.4. Statistical tests on meteorological variables exclude the possibility that the average meteorological conditions in the lockdown period of the two years 2019-2020 are different (see table 4.3 in section 4.4 for detailed results). We can therefore hypothesize that the environmental conditions were similar and that the variation in the NO_2 concentration was driven by the change in human activities. This surmise, which considers only the average conditions, is further investigated through the causal impact analysis (see the next section). Finally, the total load in the lockdown period in the year 2020 is significantly smaller than that recorded in the year 2019 (t-test on average loads difference gives a p-value $< 10^{-6}$), while there is no significant difference in the pre-lockdown periods. It is not possible to test the differences in mobility trends due to lack of publicly available data for the year 2019, but a visual analysis of fig. 4.1 makes evident the pronounced decrease in mobility during the lockdown.

4.3 Building the causal nexus

In the physics of complex systems the reconstruction of the topological network structure of a system is a known problem attracting increasing interest in many fields [239, 241, 225, 316, 317]. Here, we apply two techniques – namely the partial correlation and the Granger causality – to reconstruct the network of statistical and causal relations between the components of the socio-environmental system. In fig. 4.3 are reported the structures of the two resulting networks (see also fig. 4.4 and the section 4.4 for details on partial correlation and Granger tests). The partial correlation network in fig. 4.3a) shows the statistical relation between pairs of system components, after removing the effect of the other components. This structure reveals the stronger relations between all human activity variables, which vary synchronously, driven by the lockdown interventions. In particular the variable “change in movement" is the most connected. The negative relation between “residential" and “work-places" is well captured by the method, likewise for the “total precipitation", “surface radiation" and “temperature". The NO_2 variable is related just to the solar radiation, which may be interpreted as the influence of the seasonal effects mentioned in the introduction.

Figure 4.3b) shows the Granger causality network that represents the (Granger) causal influence of a variable onto another (see *Appendix* at the end of this chapter for more details on the definition). This network, that is directed and appears denser than the previous one, points out the influence of the human activities on

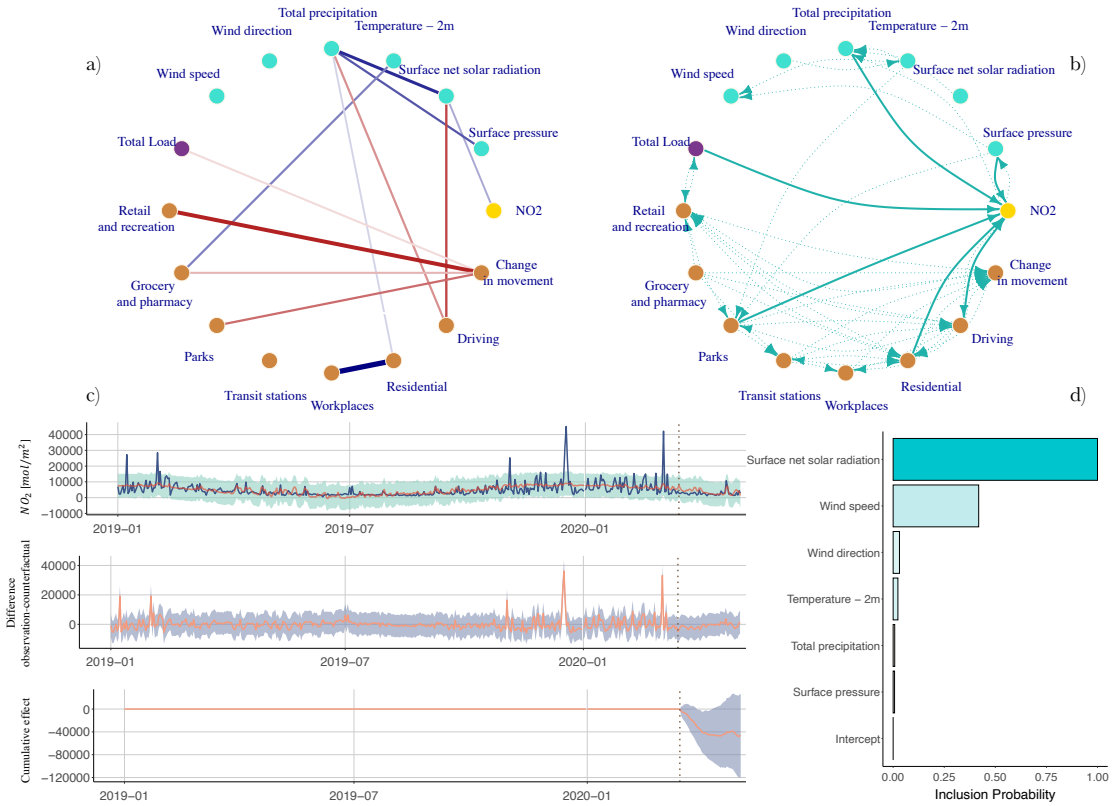


Figure 4.3 – Causal analysis for the time course of the 16 observables – a) Partial correlation network: each node corresponds to a variable, the color encodes the type of variable (meteorology, energy, mobility, NO_2); blue edges represent the negative partial correlation, while red edges represent positive partial correlation; the thickness of the edges is proportional to their partial correlation value. **b)** Granger causality network: the arrows are oriented in the causal direction; the variable which have a causal impact on NO_2 are better highlighted with solid edges. **c)** Bayesian state-space model: The top panel shows the data and a counterfactual prediction for the lockdown period. The middle plot shows the difference between observed data and counterfactual predictions. The bottom plot is the cumulative effect of the lockdown. **d)** Probability of inclusion for the regressors; light-blue bars represents negative coefficients, while red bars represents positive coefficients. Note that only the meteorological regressors can be used for the counterfactual prediction, since they are the only variables not influenced by the lockdown intervention.

the NO_2 concentration and also the possible influence of meteorological conditions that could cause variation on air pollution, such as precipitation. Meaningful meteorological relations can be found in this network, whereas the human activities constitute a dense separated cluster.

On the one hand, the partial correlation and the Granger causality reveal the complex network of interaction between the variables of the system. The topology of the interaction strongly impacts the collective dynamics of the variables, and is

essential to understand the function of the entire system. On the other hand, for the evaluation of the impact that a large scale "intervention" – as the lockdown restrictions – may have on the socio-environmental system, we need to assess the causal relation between the specific event of the lockdown, and the measured effects on the variables under study. In particular, we want to evaluate the causal relation between the concentration of pollutants and the lockdown intervention. To this aim we employed a Bayesian structural time-series model to build a counterfactual prediction of what would have happened to the NO_2 concentration if the lockdown were not implemented (see the *Appendix* at the end of this chapter for details).

The results are shown in fig. 4.3 c-d). The plot at the top of fig. 4.3c) shows the NO_2 concentration data and the counterfactual prediction for the lockdown period. The NO_2 concentration during the lockdown period had an average value of $2.76 \cdot 10^3$ mol/m^2, whereas in the absence of an intervention, we would have expected an average response of $3.71 \cdot 10^3$ mol/m^2. The difference between NO_2 concentration data and the corresponding counterfactual prediction (middle plot in fig. 4.3c)) yields an estimate of the causal effect of the lockdown. This effect is $-0.96 \cdot 10^3$ mol/m^2 with a 95% interval of $[-2.55 \cdot 10^3, 0.57 \cdot 10^3]$. In relative terms, the response variable showed a decrease of -26% with a 95% interval of $[-69\%, +15\%]$. To obtain the cumulative impact of the lockdown, the concentration data are summed up (bottom plot in fig. 4.3c)) obtaining a cumulative concentration of $132.37 \cdot 10^3$ mol/m^2 which would have been equal to $178.31 \cdot 10^3$ mol/m^2 in absence of the lockdown. Even though the results seems to reveal a clear causal impact of the lockdown on the NO_2 concentration, the computed probability of obtaining this effect by chance is $p = 0.114$. This is an evident signal of lack of statistical significance. To sum up, considering also the results of the statistical tests described above, the reduction in the NO_2 concentration are significant, but this reduction may not be caused just by curbed human activities. This fact, although restricted to the specific case study, may have relevant implications from both scientific and policy-making standpoints discussed in the following section.

4.4 Conclusions and outlook

To understand the effects of large-scale human mobility on the environment, the social and environmental systems cannot be regarded as "closed systems". In fact, the collective behaviour of the socio-environmental system depends on the mutual feedbacks between these two counterparts, that in turn give rise to emergent properties of different nature [303]. To recognise such feedbacks, the physics of complex systems leverages on the characterization of the structure of relations between the dynamical components of a system. Capitalizing on powerful tools from network science introduced in chapter 1, we revealed the structure of functional relations

between the components of a socio-environmental system, here considered as a complex network. Moreover, by coupling data-analysis and causal inference we studied the impact of a large-scale intervention – specifically a forced restriction on human mobility – on the socio-environmental system, identifying the shifts in its behaviour and providing a formal identification of the related causes.

To this aim we have fused heterogeneous data sources – including human mobility, total energy load, meteorological conditions and NO_2 concentrations – in four different periods over 2019 and 2020. These variables are considered as signals coming from the different components of the socio-environmental system. From this information, we determined the impact of the lockdown intervention on the components detecting a regime shift and assessing the significance of any different conditions by means of statistical tests. In particular, we designed and applied an information theoretic technique to find the date of a possible regime shift in the data which turns out to be few days after the imposition of the lockdown. We have found evidence that, concomitantly with the reduction in both the mobility and the energy demand, the NO2 average concentration significantly decreases in 2020 lockdown w.r.t the same period in the previous year. The lower variance of NO_2 during the 2020 lockdown is even more visible than the decrease in average concentration. This is attributable to the complex nature of the human-environmental system where internal factors may act as filter on the rapid variations of the NO_2. Shedding light on this factor would unveil new possible features of the system under consideration.

The functional organization of the complex system was reconstructed from the data using partial correlation and Granger causality. The results provide a topological map of the socio-environmental system, showing the mutual influence of the variables. Finally, we investigated the effects of the lockdown intervention relying on a counterfactual Bayesian model. Overall, the analysis detected the sign of a causal relation between the relaxation of a broad spectrum of human mobility and energy consumption and air pollution abatement during the lockdown in Northern Italy. However, the statistical significance of the causal impact result is questionable; this has important implication from both statistical and policy-making perspectives: the lack of significance suggests that the effect might be due to chance with non negligible probability, but this can also be due to other uncontrollable issues related to the data, and it cannot be undoubtedly interpreted as a lack of causal relation. For example, although the meteorological-seasonal variables, which are not affected by the lockdown intervention and are related to the NO_2 concentrations, are good candidates as explanatory variables of the NO_2 in the Bayesian model, they may still not have sufficient predictive power on the NO_2. In fact, the ideal experimental setup should involve a stronger driver of NO_2 concentration as an additional explanatory variable, which must not be

affected by lockdown interventions to produce the counterfactual prediction (as further explained in the *Appendix* at the end of this chapter). However, it is difficult to identify such a variable, since the main sources of NO_2 are anthropogenic, and virtually any anthropogenic source of pollutant was affected by the lockdown, making these variables unusable in the causal model. From the environmental policy-making perspective, the lack of a clear causal relation between the lockdown and the NO_2 concentration reduction should point out that reducing emissions from mobility and power plants may be not as effective as previously thought. For example, heating and cooking systems, as well as industrial and agricultural emissions should be better monitored and controlled. Nevertheless, although our approach is general, our result is case-specific and its robustness should be further assessed by performing analyses on a broader geographical region. In fact, as we showed in previous sections, the particular environmental conditions (e.g. meteorology, topography, etc.) have a decisive impact on pollutant concentrations. For this reason, policies should be designed assuming a systemic point of view while keeping a context-specific insight.

In conclusion, our results suggest that limiting mobility and non-essential human activities may reduce the pollution concentration, but it may not be a resolutive nor definitive remedy. The weak causal impact of the intervention can be the sign of the predominant role played by the emissions from essential and unstoppable human activities. Furthermore, other external conditions (e.g. meteorology, topography, etc.) may play antagonistically, undermining the positive effects of the restrictions. In addition, living in lockdown conditions is proving to be economically unsustainable [318]. In the light of these findings, we consider our approach to be indicative, but not definitive, for investigating the nexus between human activity and environmental conditions during COVID-19 lockdown in Italy. Further developments should complement our analysis with mobility data of 2019 and including more detailed data such as stratified mobility (i.e. heavy and light transport with differentiated emissions).

We firmly believe that a paradigm shift towards a systemic view is necessary, and fundamental, when studying the complex relations between society and environment. We argue that the synergy of the human mobility system and the environmental conditions may induce long-term adaptations of the socio-environmental system, triggering feedback loops that may alter the mobility patterns is response to environmental changes and vice versa. For this reason, the combination of causal models and complex systems science will enrich the debate on the coupled human-environmental systems. Such an approach would also be of tremendous help for decision making processes allowing for more informed and integrated choices, especially for the development of mitigation policies in accordance with climatic and environmental goals (e.g. Sustainable Development Goals of Agenda 2030).

Variable	Average				Std. Deviation			
	Group 1.1	*Group 1.2*	*Group 2.1*	*Group 2.2*	*Group 1.1*	*Group 1.2*	*Group 2.1*	*Group 2.2*
NO2 [mcg/m2]	6367.59	3747.44	6891.16	2677.80	2822.47	1618.58	6853.91	1447.24
Surface pressure [Pa]	95227.33	94666.43	94729.76	94938.83	700.79	763.56	853.13	558.27
Surface solar rad. [kJ/m^2]	5108.60	7817.19	4632.62	8142.56	1244.45	2144.71	1182.88	2335.69
Temperature - 2m [K]	277.45	281.95	277.64	282.32	2.21	2.06	1.94	3.64
Total precipitation [mm]	0.42	2.35	0.92	1.20	1.19	3.82	2.11	2.15
Wind direction [°]	-48.73	-32.45	-51.94	-7.52	25.84	55.36	27.21	52.37
Wind speed [m/s]	1.06	1.03	1.04	0.95	0.53	0.50	0.57	0.55
Total Load [MWh]	37.60	33.87	37.05	27.86	4.20	4.98	4.04	3.13
Retail and recreation [%]	-	-	-23.00	-85.98	-	-	22.08	6.19
Grocery and pharmacy [%]	-	-	-2.82	-50.23	-	-	11.85	15.82
Parks [%]	-	-	1.96	-77.31	-	-	32.62	10.52
Transit stations [%]	-	-	-30.21	-82.94	-	-	25.26	4.04
Workplaces [%]	-	-	-18.32	-67.63	-	-	16.04	7.79
Residential [%]	-	-	8.79	31.77	-	-	8.58	5.72
Driving [%]	-	-	-8.93	-80.25	-	-	33.33	5.08
Change in movement [%]	-	-	-27.09	-66.04	-	-	13.73	7.82

Table 4.1 – Main descriptive statistics of the data items divided by time groups.

Appendix

Overview of the data sets

In this chapter we relied on data of nitrogen dioxide (NO_2 concentrations) from Copernicus Sentinel-5P satellite from the 1st of January 2019 to the 1st of June 2020 (TROPOMI Level 2 Nitrogen Dioxide total column products. Version 01. European Space Agency [319]). In particular, we referred to high-resolution daily concentrations of the tropospheric NO_2 over Lombardia region. Data of meteorological conditions are retrieved from the Copernicus Climate Change Service [320] (ERA5-Land reanalysis dataset) and consist of hourly data of 6 variables – i.e surface pressure, surface net solar radiation, temperature, total precipitation, wind direction and wind speed – over the Lombardia region. These data were averaged daily. Google mobility data [306] are provided in terms of daily length of stay at different places – e.g. residence, grocery, parks, etc. – aggregated at regional level. Apple data [307] are provided in terms of variation in the volume of *driving* directions requests while Facebook data [308] in terms of positive or negative *change in movement* relative to baseline (February 2020). As for NO_2, we considered the case of the Lombardia region covering collectively the time period from 13th of January to the 27th of July 2020. All these mobility data are available only for the year 2020. For what concern energy load, we considered data of Northern Italy for the period from 1st of January 2019 to the 27th of July 2020 from the Italian transmission system operator *Terna* [321]. "Northern Italy" is the smallest available space aggregation which includes Lombardia; since Lombardia has the highest energy demand in this area, Northern Italy data are deemed appropriate for our purposes. In table 4.1 are reported the main descriptive statistics of all the

variables.

Statistical and Causal Analysis

To rigorously assess the differences in the time series of the considered variables, we tested the differences in mean between each time group with t-tests and surrogate data tests. In addition, we evaluated the differences in the variances between the groups with Fligner-Killeen test, that is a non-parametric test for homogeneity of group variances based on ranks, robust against non-normal data [312].

The magnitude of the differences observed in the time series are evaluated through two effect-size measures: the Cliff-δ [314] and the C.L.E.S. (Common Language Effect Size) [315]. The Cliff-δ is computed by enumerating the number of occurrences of an observation from one group e.g. x_{i1} having a higher response value than an observation from the second group (e.g. observation x_{j2}), and the number of occurrences of the reverse:

$$\delta = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \operatorname{sgn}(x_{i1} - x_{j2})}{n_1 \times n_2} \tag{4.3}$$

where

$$\operatorname{sgn} x := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{4.4}$$

the indices 1 and 2 refer to the two groups under consideration (as defined previously in this section) and the two time series x_{i1} and x_{j2} are of size n_1 and n_2 respectively. This statistic measures the tendency of each x_{i1} in group 1 to be higher than each x_{j2}, in group 2, and it is not dependent on any assumptions whatsoever. The values of the statistic can run from -1.0 (nonoverlapping distributions with smaller x_{i1}) to 1.0 (nonoverlapping with smaller x_{j2}). Similarly, the C.L.E.S. is defined as the probability that a randomly selected individual from one group have a higher score on a variable than a randomly selected individual from another group. This measure has the advantage of being simply interpreted as the probability of a value from one time series to be higher than a value in the other [315]. We computed this measure numerically with a brute-force approach, by random sampling repeatedly the values from the time series of the two years and computing the frequency with which the values of the first series were higher than the ones in the other.

Afterwards, to obtain a statistical indication of the possible causal relations between the 16 variables, we build the partial correlation matrix. The reconstruction of network structures using correlation measures is a very well-known technique in network science (see e.g. [239, 235, 12]). The partial correlation measures the

strength and the direction of the (rank) dependence of two variables from a set of random variables when the influence of the remaining variables is removed. More precisely, we computed the partial correlation coefficient for each pair of variables X and Y with the effects of the remaining variables $\mathbf{Z}$ removed. This can be done by performing two separate linear regressions on the variables X and Y using the vector $\mathbf{Z}$ as a regressor:

$$\begin{aligned} \mathbf{w}_x^* &= \arg\min_{\mathbf{w}} \left\{ \sum_{i=1}^N \left(x_i - \langle \mathbf{w}, \mathbf{z}_i \rangle\right)^2 \right\} \\ \mathbf{w}_y^* &= \arg\min_{\mathbf{w}} \left\{ \sum_{i=1}^N \left(y_i - \langle \mathbf{w}, \mathbf{z}_i \rangle\right)^2 \right\} \end{aligned} \tag{4.5}$$

where $\mathbf{w}$ is the vector of the regression coefficient and $\mathbf{w}^*$ is its optimal OLS estimation. After calculating the residuals $e_{x,i}$ and $e_{y,i}$ for the two variables as

$$\begin{aligned} e_{x,i} &= x_i - \langle \mathbf{w}_x^*, \mathbf{z}_i \rangle \\ e_{y,i} &= y_i - \left\langle \mathbf{w}_y^*, \mathbf{z}_i \right\rangle \end{aligned} \tag{4.6}$$

the partial correlation coefficient is computed as the Pearson correlation between the residuals:

$$\boldsymbol{r}_{XY|Z} = \frac{N \sum_{i=1}^N e_{x,i} e_{y,i}}{\sqrt{N \sum_{i=1}^N e_{x,i}^2} \sqrt{N \sum_{i=1}^N e_{y,i}^2}} \tag{4.7}$$

Subsequently, we tested the significance of the obtained partial correlation values comparing them to the results from a surrogate analysis, consisting in the estimation of the partial correlations using numerous reshuffled (not correlated) versions of the original time series [322]. Therefore, only the statistically significant relations were considered in the reconstructed network.

We further investigated the structural features of the system, studying the network of causal dependence of the variables through the “predictive causality" measure by C. Granger. The Granger causality is a tool of increasing interest in the study of the structural features of complex systems [181, 323, 324, 325, 326]. It is defined as a causality test between two time series X and Y, which detect the ability of one variable to predict the other. In other words, it says that if the information about the trajectory of Y improve the prediction of the trajectory of X, then Y causes X. Specifically, an autoregressive model is estimated for the variables X and Y in the form

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} A_{xx}^i & A_{xy}^i \\ A_{yx}^i & A_{yy}^i \end{pmatrix} \begin{pmatrix} X_{t-i} \\ Y_{t-i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{pmatrix} \tag{4.8}$$

where i is the lag, and the coefficient A^i of the model are optimized (max-likelihood). Then, if $A_{xy}^i \neq 0 \quad \forall i$ then it can be concluded that $\boldsymbol{Y}$ «Granger causes» $\boldsymbol{X}$. To

obtain a statistically significant conclusion, a Wald test [327] is performed which evaluates the null hypothesis $H_0 : A^i_{xy} = 0$. It is to be noted that, since in general $A_{xy} \neq A_{yx}$, also the direction of the causal relation can be assigned, allowing to distinguish the situation in which $\boldsymbol{Y}$ «Granger causes» $\boldsymbol{X}$ from $\boldsymbol{X}$ «Granger causes» $\boldsymbol{Y}$; consequently, the network will be directed as well. Since the Granger test relies on an autoregressive model, the time series should be stationary. In our case, the observed time series are not stationary – according to the Augmented Dickey–Fuller test. Therefore, we differentiated the data obtaining a set of time series that passed the unit root test.

Finally, to assess the impact of the lockdown on the pollutant concentration we employed a Bayesian modeling technique for casual inference which uses a structural time-series model to predict what would have been the system evolution after an intervention, if the intervention had never occurred [165]. Structural time-series models are state-space models for time-series data. They can be generally defined as follows:

$$y_t = Z_t^{\mathrm{T}} \alpha_t + \varepsilon_t \tag{4.9}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \tag{4.10}$$

where $\varepsilon_t \sim \mathcal{N}(0, \sigma_t^2)$ and $\eta_t \sim \mathcal{N}(0, Q_t)$ are independent of all other unknowns. Equation (4.9) serves as a link between the observed data y_t and the latent state vector α_t which is a d-dimensional vector. In eq. (4.10) is described the evolution of the latent state α_t. The term Z_t is a d-dimensional output vector, T_t is a $d \times d$ transition matrix, R_t is a $d \times q$ control matrix, ε_t is a scalar observation error with noise variance σ_t, and η_t is a q-dimensional system error with a $q \times q$ state-diffusion matrix Q_t where $q \leq d$. In our case the observation term y_t is the vector of observation of NO_2. The general structure of eq. (4.9) and eq. (4.10) can be adapted to describe different behaviours of the latent state, including local trends, seasonality and the influence of covariates $\mathbf{x}_t$ (for more details on the model structure see e.g. [165, 164]). The covariates, which are important actors for the aim of our analysis, can be included in the model considering a static regression by setting $Z_t = \beta^1 \mathbf{x}_t$ and $\alpha_t = 1$, or by dynamical regression

$$\mathbf{x}_t^{\mathrm{T}} \beta_t = \sum_{j=1}^{J} x_{j,t} \beta_{j,t} \tag{4.11}$$

$$\beta_{j,t+1} = \beta_{j,t} + \eta_{\beta,j,t} \tag{4.12}$$

where $j = 1, \ldots, J$, are the index for each covariate, $\eta_{\beta,j,t} \sim \mathcal{N}\left(0, \sigma^2_{\beta_j}\right)$, $\beta_{j,t}$ is the coefficient for the j-th control series and σ_{β_j} is the standard deviation of its associated random walk. This regression component can be turned into the same

structure as eq. (4.9) and eq. (4.10) by setting $Z_t = \mathbf{x}_t$ and $\alpha_t = \beta_t$ and by setting the corresponding part of the transition matrix to $T_t = I_{J\times J}$, with $Q_t = \text{diag}\left(\sigma^2_{\beta_j}\right)$. In our case, we used the static regression in order to prevent overfitting [328].

The parameter estimation as well as the model simulation are conducted in a Bayesian framework, so that empirical priors can be incorporated on the model parameters [165]. Let θ generically denote the set of all model parameters and let $\alpha = (\alpha_1, \ldots, \alpha_m)$ denote the full state sequence. A prior distribution $p(\theta)$ on the model parameters as well as a distribution $p(\alpha_0 \mid \theta)$ on the initial state values are needed to define the model. The parameters of the model in eq. (4.9) and eq. (4.10) are the set of variances, for which a commonly used prior distribution is the *Gamma* distribution $\mathcal{G}(a, b)$ with expectation a/b:

$$\frac{1}{\sigma^2} \sim \mathcal{G}\left(\frac{v}{2}, \frac{s}{2}\right). \tag{4.13}$$

The hyperparameters can be interpreted as a prior sum of squares s, so that s/v is a prior estimate of σ^2, and v is the weight, in units of prior sample size, assigned to the prior estimate. The values of α and θ can be sampled from $p(\alpha, \theta \mid \mathbf{y})$ using Markov Chain Monte Carlo, subsequently the counterfactual time series $\tilde{\mathbf{y}}_{n+1:m}$ are sampled from the predictive posterior distribution $p(\tilde{\mathbf{y}}_{n+1:m} \mid \mathbf{y}_{1:n})$ (see [165] for more details on the model estimation).

For this method to work correctly, the covariates themselves must not be affected by the intervention and the relationship between covariates and treated time series, as established during the pre-intervention, must remains stable throughout the post-intervention period. Specifically, this means that the meteorological/seasonal variables used as covariates are considered to be not affected by the lockdown intervention, which seems a reasonable assumption. Given this assumption and considering that the relation between meteorological conditions and NO_2 concentration remain stable after the lockdown, a significant deviation from the counterfactual NO_2 concentration would indicate the causal impact of the intervention. This means that the lockdown would be considered as the event causing the drop in air pollution through activity restrictions. On the contrary, if the main drivers of the NO_2 concentration were the meteorological/seasonal conditions (and not human activity), we would expect the data to follow the counterfactual prediction in the post-treatment period. In that case, no significant causal impact of the intervention would be found.

Thanks to this approach we stated the impact of the reduced human mobility and energy consumption due to the lockdown, considered as the attributable intervention on NO_2 concentrations.

Statistical tests results

	Test p-values		Effect Size	
Groups	*t-test*	*Fligner test*	*Cliff δ*	*C.L.E.S.*
Groups 1.1-2.1	0.53	0.28	0.01	0.51
Groups 1.1-1.2	$5.80 \cdot 10^{-6}$	$1.38 \cdot 10^{-4}$	0.60	0.80
Groups 1.1-2.2	$< 10^{-8}$	$1.40 \cdot 10^{-}6$	0.82	0.91
Groups 1.2-2.1	0.99	$2.43 \cdot 10^{-5}$	-0.56	0.22
Groups 2.1-2.2	$< 10^{-8}$	$< 10^{-8}$	0.79	0.89
Groups 1.2-2.2	$7.39 \cdot 10^{-4}$	$3.51 \cdot 10^{-7}$	0.72	0.81

Table 4.2 – Complete results of statistical tests and effect size (related to the t-test) comparing different groups for the variable NO_2.

	Wilcoxon test p-values					
Variables	*Group 1.1-2.1*	*Group 1.1-1.2*	*Group 1.1-2.2*	*Group 1.2-2.1*	*Group 2.1-2.2*	*Group 1.2-2.2*
Surface pressure	0.022	0.001	0.066	0.478	0.422	0.032
Surface solar rad.	0.061	$< 10^{-8}$	$< 10^{-8}$	$< 10^{-8}$	$< 10^{-8}$	0.354
Temperature -2m	0.964	$< 10^{-8}$	$< 10^{-8}$	$< 10^{-8}$	$< 10^{-8}$	0.184
Total precipitation	0.004	$2.32 \cdot 10^{-5}$	0.022	0.030	0.789	0.046
Wind direction	0.659	0.147	$9.6 \cdot 10^{-6}$	0.106	$4.20 \cdot 10^{-6}$	0.015
Wind speed	0.359	0.945	0.720	0.825	0.366	0.252

Table 4.3 – Complete results of statistical tests comparing average values of the meteorological variables in different periods. It is to be noted that during the pre-lockdown periods (1.1-2.1) only the precipitations were significantly different ($\alpha = 0.01$), while during the lockdown periods (1.2-2.2) all the meteorological conditions were not significantly different.

(a) Partial correlation test p-values **(b) Granger test p-values**

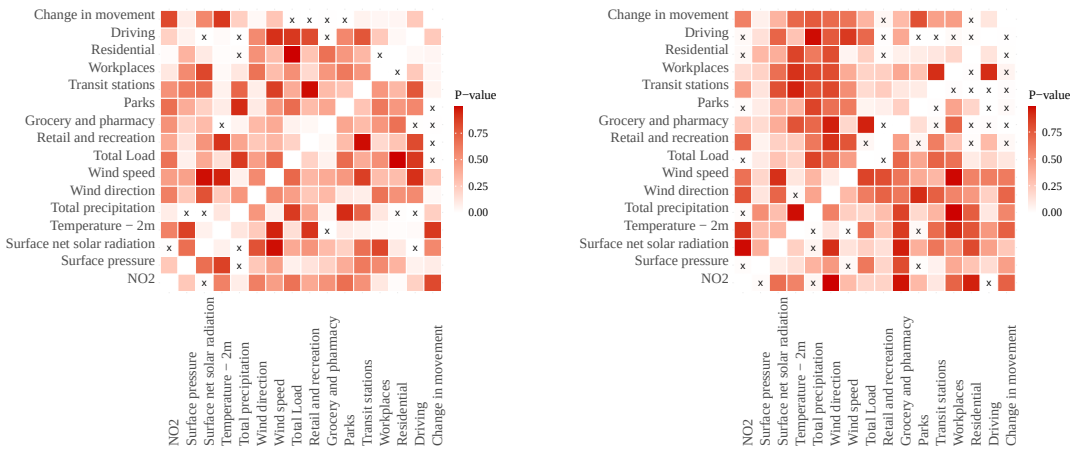


Figure 4.4 – **Connectivity matrices (p-values)** – Results of partial correlation and Granger tests. The color of the entries indicates the p-value estimates. Black crosses are placed where the null hypothesis (no relation between variable) is rejected. The significant relations are depicted in the networks of fig. 4.3a) and b).

Chapter 5

The causes: human displacement caused by environmental disasters

In brief

We present an analysis of a dataset of human mobility kindly provided by Facebook in 2019. The data are about the variation in mobility fluxes due to various environmental disasters. Here we analyse two case studies, regarding tropical cyclones that afflicted Southern USA and Philippines in 2020. We show the change in mobility due to the disaster, and we estimate a gravity model using ridge regression. Our results about the significant predictors of mobility flows are in accordance with scientific literature. Moreover, the recently introduced Social Connectedness Index proves to be associated with higher mobility flows.

5.1 Introduction

> Unfortunately, we are a species with schizoid tendencies, and like an old lady who has to share her house with a growing and destructive group of teenagers, Gaia grows angry, and if they do not mend their ways she will evict them.
>
> *The Revenge of Gaia*, James Lovelock

The relation between environmental changes and human mobility is currently a central topic in various research fields. Many factors are involved in determining

mobility patterns, as we sketched in chapter 1. Environmental change can be a major *cause* of migration and people displacements, and the lack of data has played against the understanding of this phenomenon. This is particularly problematic when people are displaced by catastrophic environmental events. The traditional empirical data collection, such as census data, local surveys, tax revenue data, are often unable to trace mobility in this situations. Nevertheless, understanding and predicting mobility fluxes before, during and after disasters is key to organize and activate effective humanitarian operations and to support long-term planing for reconstruction and adaptation. In this regard, Facebook have recently released the "Disaster Maps" dataset, with the express purpose of "helping organizations address the critical gap in information they often face when responding to natural disasterss" [329]. This dataset contains geolocated mobility fluxes between locations during disastrous environmental events [330]. Unfortunately, at the time when the following analysis was conducted, the dataset was still incomplete, and a comprehensive analysis had to be deferred.

In this concise chapter we present a preliminary analysis of the Facebook Disaster Maps for a couple of case studies, that reveal some elements of the relationship between environmental system and mobility patterns. In the previous chapter, we showed how the environmental system responds to large-scale variations of human mobility, conversely, here we deal with the response of human mobility system to an abrupt environmental change. The focus is on shot-term and regional mobility variation following ,

5.2 Data and pre-processing

For our analysis we combined the Facebook Disaster Maps dataset [330] and hurricanes trajectories from the "International Best Track Archive for Climate Stewardship" (IBTrACS) dataset from the NOAA [331, 332]. The Disaster Maps consist of geolocated information about where individuals are located, how they are moving, and where they are checking in safe during a natural disaster [329]. The data are de-identified and spatially aggregated to a 360'000m^2 tile or local administrative boundaries. The movements are measured during a baseline period (bilateral movements averaged across the three weeks prior to the disaster) as well as during the event, in order to determine the percentage change in mobility flows caused by the event. Besides the flows, the population at the locations of origin and destination are also given. The IBTrACS data merges tropical cyclone storm track datasets from agencies around the world to create a global database. The information include the geographical coordinates of the cyclones, their Maximum Sustained Wind (MSW) speed, the atmospheric pressure and other meteorological variables. To control for the socio-economical drivers of mobility, we employed

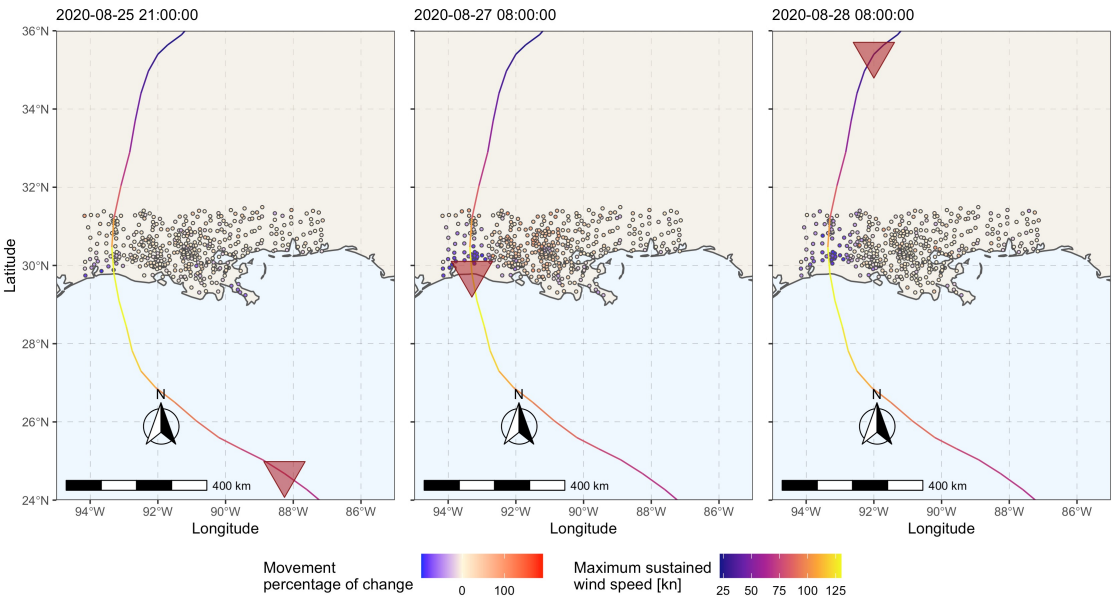


Figure 5.1 – Disaster Map – Changes in human mobility before (left), during (center) and after (right) the passage of the hurricane *Laura* in Southern USA in 2020. Each point corresponds to a measurement unit of the Disaster Map and its color encode the percentage change in human mobility with respect to the baseline. The red triangle represents the eye of the storm.

the spatially distributed GDP (PPP) and the Human Development Index (HDI) from [333]. These data are gridded over the whole world at 5 arc-min resolution for the 25-year period of 1990–2015. The total GDP (PPP) is also provided with 30 arc-sec resolution for three time steps (1990, 2000, 2015). We used the data for the year 2015. In addition, we used the Social Connectedness Index (SCI) by Facebook [334, 335], recently proposed as a measure of the social relations of individuals from different areas.

The data have been initially processed using a GIS software, to define the attributes of the areas matching the locations of origin and destination of the Disaster Maps. For each site we got the related population, GDP (PPP), HDI and orthogonal geodesic distance from the trajectory of the cyclone. For each pair of locations we have also the SCI and we computed the relative distance.

5.3 Results and Methods

Figure 5.1 presents an example of the changes in mobility fluxes at three time steps in the course of the hurricane Laura in Alabama and neighbouring states, in 2020. The figure shows an increase in movements at the passage of the hurricane in the areas around the center of the storm, while they are strongly reduced closer

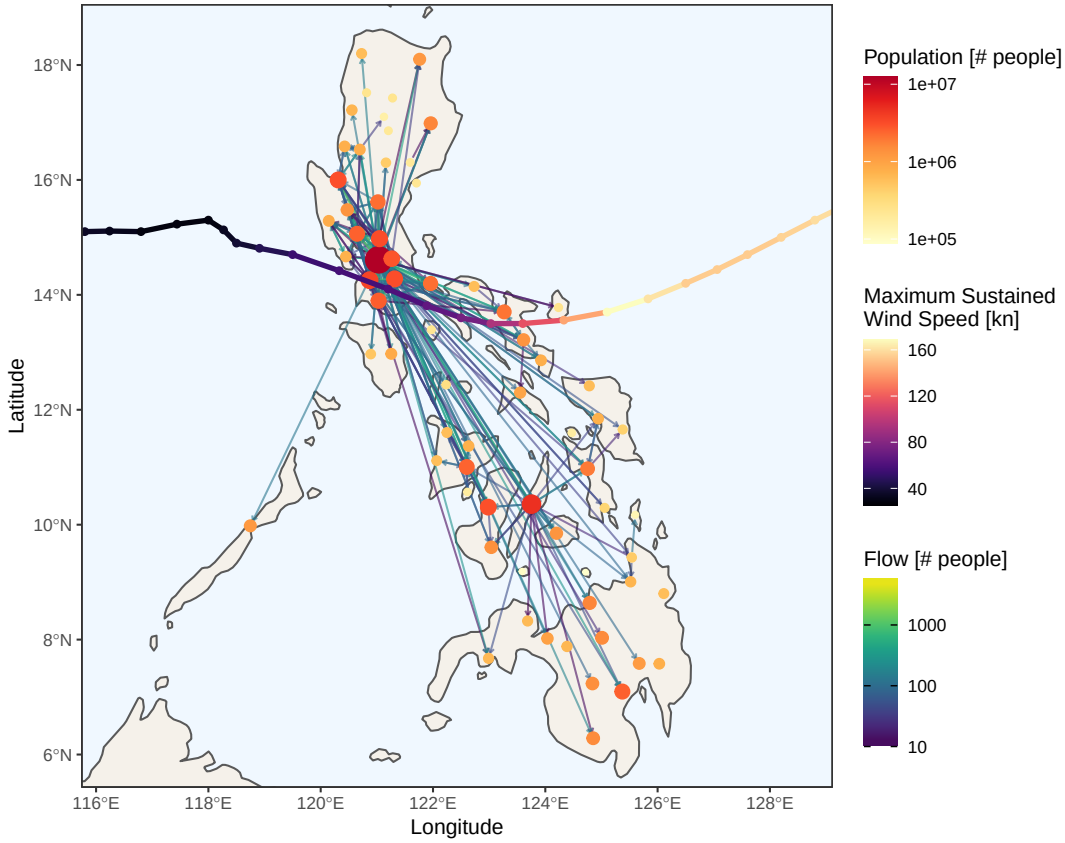


Figure 5.2 – Mobility network of displaced people after typhoon Goni – The nodes of the network on the map represent the locations of origin and destination, the edges are the flows of people, aggregated over the two weeks after the landfall of the typhoon. The trajectory of the typhoon goes from left to right and is colored according to the MSW speed.

to the landing zone and on the trajectory of the hurricane eye, continuing to decline also after the passage. Rapid-onset disasters of this kind have usually little impact on large-scale human movement, that are mostly limited to temporary and short-range movement or immobility [80].

A similar example comes from Philippines, during the passage of the Goni typhoon, one of the strongest tropical cyclone in world history, a Category-5 Saffir–Simpson scale typhoon. This event caused the death of at least 20 people and displaced 400,000 people in Bicol region, when it made landfall over the Philippines on November 1st 2020. The trajectory of the typhoon and the mobility network is represented in fig. 5.2. The typhoon passed south-west of Manila, the capital city of Philippines and the main source of mobility. While wealthier communities like Manila, have better access to evacuation centers and to resources to rebuild, in many rural communities, housing conditions is often poor, and loss of housing led

to increased overcrowding in shared and communal shelters, increasing the risk of COVID-19 transmission [336].

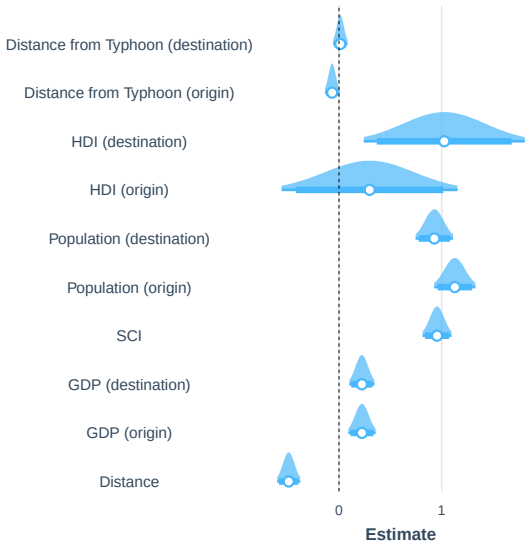
Using the data described in the previous section, we conducted a preliminary analysis by estimating a gravity model (described in details in section 2.1), using the following predictors:

- **Population** at locations of origin and destination
- **Gross Domestic Product** based on Purchasing Power Parity (GDP-PPP) at both location of origin and destination;
- **Orthogonal geodesic distance** from the trajectory of the typhoon to each location of origin and destination;
- **Human Development Index (HDI)** is a summary measure of average achievement in key dimensions of human development, at origin and destination;
- **Social Connectedness Index (SCI)**, measures the strength of connectedness between two geographic areas as represented by Facebook friendship ties;
- **Distance** between location of origin and destination.

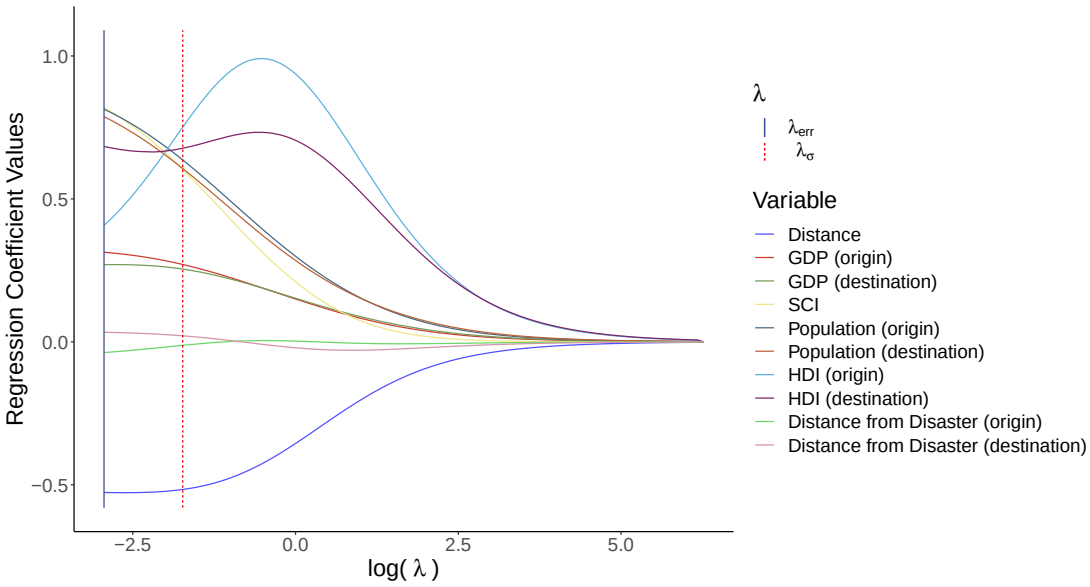
These input variables have been standardized before estimating the model parameters. In fig. 5.3a are reported the distributions of the regression coefficients, where the significance of the predictors can be assessed visually: if the light-blue colored part of the distributions crosses the 0 axes, the coefficient is not significant at significance level $\alpha = 0.05$. In compliance with other similar studies (see e.g. [121, 337]), and with the economic theories of migration, the main drivers of mobility prove to be the populations at both locations of origin and destination, the GDP and the distance between locations. Also the HDI at destination is a good predictor of the flows, as well as the Facebook Social Connectedness Index. The distance of the origin from the typhoon is (slightly) negatively correlated with mobility fluxes, and the distance of the destination is not significant.

Since we have many independent variables, some of whom may be collinear (e.g. HDI and GDP are highly correlated), we also performed a ridge regression using the same model structure of the gravity model. The ridge regression is a regularization method that shrinks the values of the estimated coefficients towards zero and allow to decrease the variance, by modifying the usual cost function used in OLS [338]. In particular, the cost function becomes

$$\sum_{i=1}^{M} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{M} \left(y_i - \sum_{j=0}^{p} w_j \times x_{ij} \right)^2 + \lambda \sum_{j=0}^{p} w_j^2 \tag{5.1}$$



(a) Gravity model coefficients estimated through OLS. The white dots indicate the position of the mean estimated value. The distributions of the regression coefficients (light-blue) span the 95% confidence intervals. The darker bars below the distributions denote the 90% confidence intervals



(b) Ridge regression coefficient for the gravity model. The term λ_{err} in the legend is the minimum value of lambda that results in the smallest cross-validation error. Instead, λ_σ is the largest value of λ such that the error is within 1 standard error of the cross-validated errors for λ_{err}; it can be chosen for a simpler model in predictions (more regularized coefficients)

Figure 5.3 – Gravity model for mobility fluxes after typhoon Goni.

were the last terms represents an additional penalty equivalent to square of the magnitude of the coefficients. Minimizing the cost function in eq. (5.1) is equivalent to solve a constrained optimization problem with the Lagrange multiplier equal to the λ term, that serves as a tuning parameter to control the relative impact of the residual sum of squares and the constraint on the regression coefficient estimates. The optimal value of lambda (λ_{err}) is achieved by minimizing the cross-validation error varying λ. In fig. 5.3b are reported the results of this procedure. It can be seen that the optimal value of λ is fairly small, indicating that the optimal fit only involves a small amount of shrinkage relative to the non-regularized model. Nevertheless, the influence of the HDI is highly reduced, and results to be less than the one of population and SCI. Even with a stronger regularization (e.g. using λ_σ as defined in fig. 5.3b), the two terms remain comparable to the ones of population. The coefficient related to the node-node distance (*dark-blue* line in the figure) remains firmly below -0.5, as in the standard gravity model. The coefficients for the GDP remain as well similar to the ones estimated with the non-regularized estimation. Finally, it is to be noticed that the association between the mobility flows and the distance of the origin from the typhoon trajectory is significant, but its relative importance seems to be low.

5.4 Conclusions

The two case studies illustrate the effects of two disastrous environmental events on human mobility patterns. The passage of a tropical cyclone has a tremendous impact on human activities, and causes people to search shelter and consequently halt mobility in the period around the event. Although these rapid-onset environmental changes usually do not lead to long-lasting displacement and immobility, the Disaster Maps show significant variations in mobility. On a longer time interval, according to the gravity model, the population, the HDI and the distance between locations have a significant association with the migration flows, in agreement with the literature (see chapter 1). Also the social connectedness index may play a role; it is known from other studies that people tend to move preferably to locations in which they had significant social bonds, even in hazardous situations [63].

In both the estimation methods illustrated above for the case of Philippines, the effect of the distance from the typhoon seems to be weaker than expected; this may be due to different facts: first, the data used here refer to two weeks after the events, and no other data are available for previous periods, closer to the date of the landfall of the typhoon Goni. Thus, we can speculate that higher mobility fluxes could have happened just before the event, or few days after. Second, of the 400'000 people displaced by the typhoon, almost 350'000 of them found a shelter in evacuation centers inside the same city or region [339, 336], while the resolution

of the dataset is too coarse to detect such fluxes.

The gravity model is useful to get an overall understanding of the mobility system, but further analysis are needed to obtain reliable quantitative results about the drivers of mobility. On the other hand, the Disaster Maps dataset may be of tremendous help to explore a wide range of conditions to validate further theoretical frameworks. This dataset may be used in future research to find a universal characterization of the routes of mobility in terms of network pattern and flows during calamitous events.

Chapter 6

The causes: Features-enriched Radiation Model

In brief

In this chapter we present our Features-Enriched Radiation Model, a generalization of the Radiation Model for human mobility. The Features-Enriched Radiation Model (FERM) is a flexible mathematical model that can be applied when, besides the population, other exogenous information can be used to model the attractiveness of geographical locations. This information is encoded in the features of the locations and acts as a *cause* of large-scale human movements. Our model is based on the same physical process of the original Radiation Model, but the stochastic process is generalized in order for the features to be included as external drivers. The features can change the mobility patterns by changing the bilateral flows, while the global influence of the surrounding locations is still governed by an emission/absorption process. In the following, we demonstrate the equivalence between our model and the original Radiation Model when the locations are considered indistinguishable. We illustrate the behaviour of the FERM using "synthetic worlds" mathematically generated and by feeding the model with real-world features derived from the *human climate niche* of Italy and United States of America. The mobility patterns generated by the FERM reflect the spatial distribution of location features.
This work was presented as a contributed talk at the *Conference on Complex Systems - CCS2021*, and it is going to be submitted to a scientific journal.

6.1 Introduction

> The least movement is of importance to all nature. The entire ocean is affected by a pebble.
>
> *Pensées* - Blaise Pascal

Throughout history, billions of people around the world have moved seeking opportunities for better living conditions. Nowadays, large-scale human mobility and migration are increasingly influenced by the broader global transformations in economy, society and environment. In turn, massive human movements have an impact on the locations of origin and destination. Understanding the rules governing human movements may have important implications in many research fields, such as urban planning [340, 341, 342], socio-economy [3, 343], epidemic spreading [344, 345, 8, 346], human rights [347, 348, 349] and geopolitics [350, 351, 352] at regional and international levels. A major problem in the conceptualization of a human mobility model is to identify the causal relation between a large set of variables which influence the travellers/migrants decisions. The socio-economical, political and environmental conditions act as exogenous drivers on individual decisions [10], leading to the emergence of complex mobility patterns. One of the first attempts to give a formal definition of the forces driving migration is due to Ravenstein in 1885 and 1889 [53, 54]. As already explained in chapter 2, the author gave some general "laws of migration" based on social and economical principles. Some of the Ravenstein laws found a mathematical expression in the law of universal gravitation, in which population and economic indices work in the place of masses. A seminal work by Schneider in 1959 [139], started the long tradition of the gravity model for human migration and trade [117, 353, 354]. In the gravity model, migration fluxes are proportional to the population at the origin and destination, and inversely proportional to the distance between them. More recently, the Radiation Model (RM) has been proposed as another physically based mathematical model for human mobility [19]. This model relies on first principles and it is adherent to known findings in sociological theories of migration such as the effect of distance [355] and the intervening opportunities [137] (see chapter 2 for details). The spatial locations are modelled as sources emitting particles that have a certain probability to be absorbed by the surrounding locations. The probability of absorption depends only on the spatial distribution of the population, which is the only input required. Consequently, the model considers the population at origin and destination as the unique forcing driver of mobility, but no other exogenous information is considered explicitly. Therefore, a change in the external conditions that does not modify the population distribution would not affect the mobility fluxes. For this reason, the model cannot take into account changes

in the location-specific features (e.g. environmental conditions, restrictions, etc.) potentially crucial to predict modifications in mobility patterns. Recently, some generalization of the Radiation Model have been proposed: Simini et al. [356] derived a new form of the RM (Radiation Model with selection) in continuous space which also generalize the intervening opportunities model; Kang et al. [357] proposed a generalized version that corrects the model for spatial scales using a scaling exponent, and adding system constraints including searching direction and trip OD constraint; more recently, Alis et al. [358] substituted the population with an urbanization index that is a function of local features, whose parameters are estimated using machine learning techniques.

In our work, we present a Feature-Enriched version of the Radiation Model (FERM) which considers explicitly the exogenous drivers that *cause* the emergence of mobility patterns. Our model, while maintaining the same physical process of the original model, is able to redirect the mobility fluxes according to the attractiveness assigned to each location. The physical emission/absorption process of the original model remains unchanged, while the stochastic process that generate the network structure is modified in a way that the probability of absorption of the particles reflects the attractiveness expressed by the features of the locations. Therefore, the model is a mathematically grounded alternative to the original Radiation Model and its current generalizations, when location features are deemed important for the prediction of mobility fluxes. In this chapter, we describe the mathematical formulation of the FERM and the numerical method for the computation of the mobility network structure. The behaviour of the model is exemplified using synthetic spatial collections of locations that exhibit a self-similar distribution. Subsequently, we describe in details the behaviour of the FERM through a couple of case studies: Italy and United States. We derive the mobility flows between Italian regions and between the states of the USA, when the features of the locations are the possible future climatic conditions given by climate scenarios at 2070. The parameters used in the following scenario analysis are not optimized to fit input data, but rather, are fixed in such a way to reproduce mobility patterns reflecting the real-world spatial distribution of location features. We will prove that, assuming an optimal parametrization, the Features-Enriched Radiation Model can predict the mobility flows by exploiting exogenous information that a simple RM cannot handle.

6.2 The Radiation Model

In chapter 2 we gave a brief introduction of the Radiation Model. In this section, we will explore the mathematical model in depth, and we will highlight the issues that led us to our generalization. The Radiation Model is a relatively new model

used to describe commuting and migration fluxes that requires as input only information about the population distribution in different locations [19]. Therefore, the *push and pull factors* of migration – described in section 1.2 – are encoded in the population that is a proxy for all other possible factors. This assumption is well suited for some socio-economic variables, such as the number of job opportunities, but may not be appropriate to describe the influence of cultural, political, geographical and environmental conditions, among others [358]. These exogenous variables are not considered explicitly as driving forces for human movements; our purpose instead, is to generalize and enrich the mathematical structure of the RM in order to improve its predictive power allowing the external conditions to be taken into account.

The formal definition of the RM follows the analogy with the radiation emission and absorption processes studied in physical sciences. Each location in space is considered as a source of radiation emitting identical and independent particles. The process is summarized as follows:

1. The location i emits a particle Q associated with an *absorption threshold* z_i^Q, which encodes the tendency of the particle to be absorbed. The larger the threshold, the less the probability to be absorbed. This

2. The location j, has a certain probability to absorb the particle Q coming from i: the absorption is more likely the greater the *absorbance* z_j^Q.

3. The particle Q is absorbed by the closest location whose absorbance is greater than its absorption threshold.

4. The process is repeated many times (many particles for each location) to obtain the fluxes between all the locations.

Both the *absorption threshold* z_i^Q and the *absorbance* z_j^Q are defined as the maximum number obtained after respectively m_i and n_j random extractions from a single probability distribution[1] $F(z)$ fixed a priori and **assumed to be the same for all the locations**. The probability density function $\frac{dF(z)}{dz} = p(z)$ is called *benefit distribution* and represents the quality of the opportunities in the different locations for the travelers [354]. In the original model, the terms m_i and n_j are the population in the locations i and j respectively [19]. Therefore, to larger populations correspond larger sample size which in turn gives higher values of z^Q on average.

Assuming that $F(z)$ does not varies between locations, the probability of one emission/absorption event between any two locations can be derived analytically.

[1] In the following we will denote with $F(\cdot)$ the cumulative distribution function (CDF))

Specifically, the probability $P(1|m_i, n_j, s_{ij})$ that a particle Q emitted from location i with population m_i is absorbed in location j with population n_j, with s_{ij} representing the total population in all locations between i and j, is given by

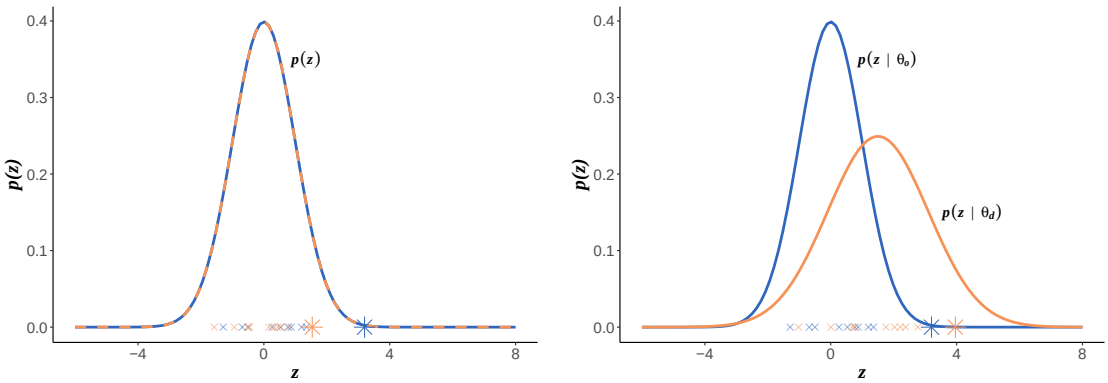
$$\begin{aligned} &P\left(1 \mid m_i, n_j, s_{ij}\right) = \\ &\int_0^\infty m_i F(z)^{m_i-1} \frac{dF(z)}{dz} F(z)^{s_{ij}} [1 - F(z)]^{n_j} dz = \\ &\frac{m_i n_j}{(m_i + s_{ij})\,(m_i + n_j + s_{ij})}. \end{aligned} \tag{6.1}$$

where the term $m_i F(z)^{m_i-1} \frac{dF(z)}{dz}$ is the probability density function of the *maximum* in m_i extractions from the benefit distribution, which gives the probability of obtaining the value z of the absorption threshold in location i. The term $F(z)^{s_{ij}}$ is the distribution function of the maximum value in s_{ij} extractions, and represents the probability that the values extracted for the locations between i and j are *always less* than the value z. Finally, the term $[1 - F(z)]^{n_j}$ is the probability that the maximum over n_j extractions is *greater* than z, i.e. the probability that the absorbance in location j of the particle Q is greater than its absorption threshold. Equation (6.1) is the mathematical formalization of the procedure listed above.

In this setting, eq. (6.1) depends only on the population m_i and n_j in locations i and j and on the population s_{ij} between them, since the particular shape of the distribution $F(z)$ is removed by integrating over all z. Consequently, the probability $F\left(1 \mid m_i, n_j, s_{ij}\right)$ is independent of the distribution $F(z)$, making the model parameter-free. Therefore, on the one hand this model is particularly simple to use and does not require any information about previous mobility flows. The only information needed is the population in each location. On the other hand, the model cannot explicitly include other relevant factors which may influence the mobility fluxes (e.g. economic and environmental conditions, social and political restrictions). Although the population may encode some information about external drivers of migration, there is no possibility to disentangle the influence of the different factors. In fact, since all the locations are characterized by the same benefit distribution $p(z)$ from which the values z_i^Q and z_j^Q are derived, the differences in the fluxes depend only on the differences in the population and in the positions of i and j relatively to the surrounding locations.

6.3 The Feature-enriched Radiation Model

In order to allow external drivers to be included in the model, we propose to enrich the Radiation Model, acting directly on the benefit distribution $p(z)$ so that it can be reshaped according to the specific features of the locations (see fig. 6.1). In this



(a) Radiation Model

(b) Features-Enriched Radiation Model

Figure 6.1 – Opportunity distribution sampling in RM and FERM – The Feature-Enriched Radiation Model acts on the benefit distributions, represented here with bell-shaped curves. The maximum of the samples from origin and destination benefit distributions determines the absorption threshold and absorbance (blue and orange asterisks respectively). In the case of the Radiation Model **(a)**, the benefit distribution is unique for all the nodes without distinction between origin from destination. Instead, the FERM **(b)** can discern more or less attractive nodes, leaving the population unchanged. The benefit distributions for the origin and for the destination can be parametrized to reflect the attractiveness of the locations. The parameters (θ in the figure) are appropriate functions of the exogenous features at each vertex. For instance, a more attractive destination will have higher absorbance (orange asterisk) with higher probability, attracting higher fluxes with respect to the case on the left.

way, we are providing a mathematically-grounded alternative to other methods and generalization based on empirical models and machine learning techniques [358, 359]. In general terms, we need to increase the probability of drawing higher value of z from locations with advantageous external conditions (e.g. higher income, favorable climatic conditions). By adapting the benefit distributions, the physical process which routes the travelers remains the same, whereas the migration flows are redistributed balancing the influence of populations and the specific features of the locations.

In our model, the benefit distribution $p(z)$, and consequently the related $F(z)$, become dependent on the features of the locations of origin and destination. Therefore, the random samples for the computation of the absorbance and absorption threshold are drawn from two distinct distributions: $F(x_o|\theta_o)$ for the origin and $F(y_d|\theta_d)$ for the destination, where θ_o and θ_d are the parameter vectors characterizing the two distributions. The size of the two samples remains equal to the population at the locations of origin and destination, and the absorption threshold and absorbance are again the maximum of the two samples. Their values are now dependent on the features of the locations of origin and destination respectively. In general, the features of an attractive location act on the parameters determining

absorption threshold and absorbance able to increase (ore decrease) the resulting probability of a flow. The probability of emission/absorption $P(1|m_i, n_j, s_{ij}, \theta_o, \theta_d)$ between two locations i and j thus becomes:

$$\begin{aligned} P\left(1 \mid m_i, n_j, L_{ij}, \theta_i, \theta_j\right) = \int_0^\infty \int_0^\infty & m_i F_X\left(x_i \mid \theta_i\right)^{m_i - 1} \frac{F_X\left(x_i \mid \theta_i\right)}{dx_i} \cdot \\ & n_j F_Y\left(y_j \mid \theta_j\right)^{n_j - 1} \frac{F_Y\left(y_j \mid \theta_j\right)}{dy_j} \cdot \\ & \prod_{k \in L_{ij}} \left[1 - F_U\left(u_{ik}\right)\right] \cdot F_U\left(u_{ij}\right) dx_i dy_j \end{aligned} \tag{6.2}$$

where X_i and Y_j are the fitness variables at origin and destination respectively, L_{ij} is the set of the indices of the locations in the circle of radius r_{ij} centered in i of length $i \to j$. More precisely, given the set of indices of all locations $\mathbf{V}$ we have that $L_{ij} = \{k \subset \mathbf{V} \mid r_{ik} < r_{ij}\ \forall k\}$. The variable U is the difference between the fitnesses: $U_{ij} = X_i - Y_j$, and is used to compute the probability of having the maximum value of X less than the maximum value of Y, since the probability of this event can be defined as $P(\bar{X} < \bar{Y}) = P(\bar{X} - \bar{Y} < 0) = F(U)$, where the the overline indicate the maximum of the variable. The first term $m_i F_X\left(x_i \mid \theta_i\right)^{m_i - 1} \frac{F_X(x_i|\theta_i)}{dx_i}$ is analogous to the first term in eq. (6.1) and represents the probability density function of the maximum in m_i extractions from $F_X(x_i)$, which gives the probability of obtaining the value $\bar{x}_i$ of the absorption threshold at the origin location i. The second term is the analogous for location of destination j. Given the independence of the two processes, the product of these two terms gives the joint probability for $\bar{X} = \bar{x}_i$ and $\bar{Y} = \bar{y}_j$. The product term $\prod_{k \in L_{ij}} \left[1 - F_U\left(u_{ik}\right)\right]$ is the probability that the values $\bar{Y}$ of the locations between i and j are *always less* than the value $\bar{X}$ extracted in i. In other words, is the probability $P(\bar{x}_i > \bar{y}_k) = 1 - F_U(u_{ik})$ extended to all $k \in L_{ij}$. The last term, $F_U\left(u_{ij}\right)$ is the probability that $\bar{x}_i < \bar{y}_j$, i.e. the probability that the absorption threshold in location j is greater than the absorbance in j.

The integral in eq. (6.2) has no simple analytical solution as in the original model. However, it is possible to compute it numerically by sampling from the distributions, to obtain an approximation of the probability of a travel between i and j $P(1|m_i, n_j, s_{ij}, \theta_o, \theta_d)$. Once the distributions $F_X(x_i \mid \theta_i)$ and $F_Y(y \mid \theta_j)$ are fixed, it is possible to sample m_i and n_j values and taking the maximum to obtain the absorbance and absorption threshold, and then compute the products in the integral. This procedure must be done several times in order to have a reliable approximation and is computationally expensive. Instead, we approach the problem in a more efficient way, by sampling a single value directly from the distribution of the maximum (DM) (or largest order statistic distribution) of the random variables X and Y. The DM can be derived analytically once the

$F_X(x_i \mid \theta_i)$ and $F_Y(y \mid \theta_j)$ distributions are known, but in general it has a complex analytical form, which requires a sophisticated sampling algorithm. We choose

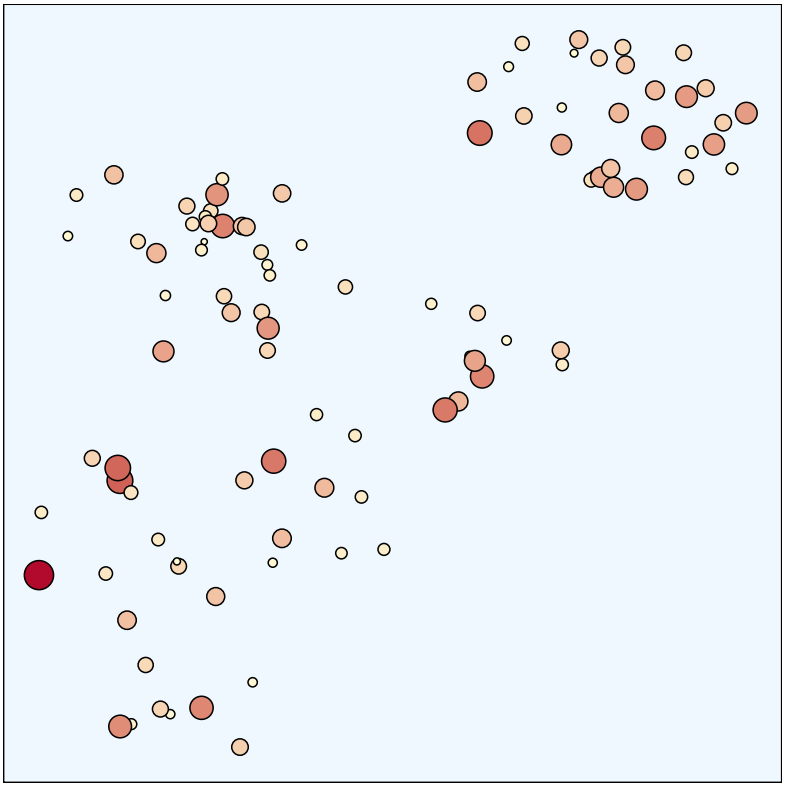


Figure 6.2 – Synthetic spatial distribution of locations – The coordinates of each location are generated using the Soneira-Peebles model. The color of the circles encode the population at each site.

the *Adaptive Rejection Metropolis Sampling* algorithm [360] to sample repeatedly from the distribution of the maximum. Subsequently, having the sampled values of the absorbance and absorption threshold, the product on the right hand side of eq. (6.2) reduces to a comparison between the two values. If the absorbance is higher than the absorption threshold, the particle Q with absorption threshold x_i will travel towards location j. Repeating this process gives the relative frequency of the travels between each pair of locations.

The actual sampled values depend on both the shape and parametrization of the benefit distribution. The more the distribution is shifted to the right, the more likely the values will be higher, reflecting the higher attractiveness of the location. In addition, flatter distributions would produce more homogeneous OD matrices, whereas picked distributions would force the flows only towards the nearest most attractive location. Given these essential properties, a natural choice for the benefit distribution is the Gaussian distribution, which is indeed the maximum-entropy

distribution with constrained mean and variance. The optimal estimation of the model parameters remains still non-trivial, requiring further efforts. In section 6.5 we provide some valid options to tackle this issue. Instead, the results of the next section 6.4 are obtained by considering a set parameters that explore different scenarios.

If the parameters are the same for all locations (i.e. $\theta_i = \theta_j \ \forall i, j$) the model in eq. (6.2) is equivalent to the standard Radiation Model in eq. (6.1). Figure 6.2 shows a realistic spatial distribution of 100 locations characterised by their population. To set the position of the nodes, we implemented a mathematical model originally designed to describe the fractal distribution of matter in the universe [361] (see the *Appendix* at the end of this chapter for more details). Each location is assigned a population sampled from a power-law distribution [362]. Figure 6.3 shows the results of a standard RM in comparison with an equivalent simulation of the FERM (with omogeneous parameters). The probability matrices and the network structures are almost identical. The FERM is able to reproduce the results of the RM with arbitrary precision, when the locations are considered equal (except fot population). This is confirmed by the very low Jaccard distance $d_{WJ} = 0.061$, which is used here to measure the difference between the two network structures. This distance is defined as $d_{JAC}\left(A_1, A_2\right) = 1 - J\left(A_1, A_2\right) = 1 - \frac{|A_1 \cap A_2|}{A_1 \cup A_2|}$, where $J\left(A_1, A_2\right)$ is the Jaccard similarity, A_1 and A_2 are the adjacency matrices, and union and intersection are meant element-wise. The Jaccard distance can be generalized to handle weighted networks, both directed and undirected: the Weighted Jaccard similarity is defined as:

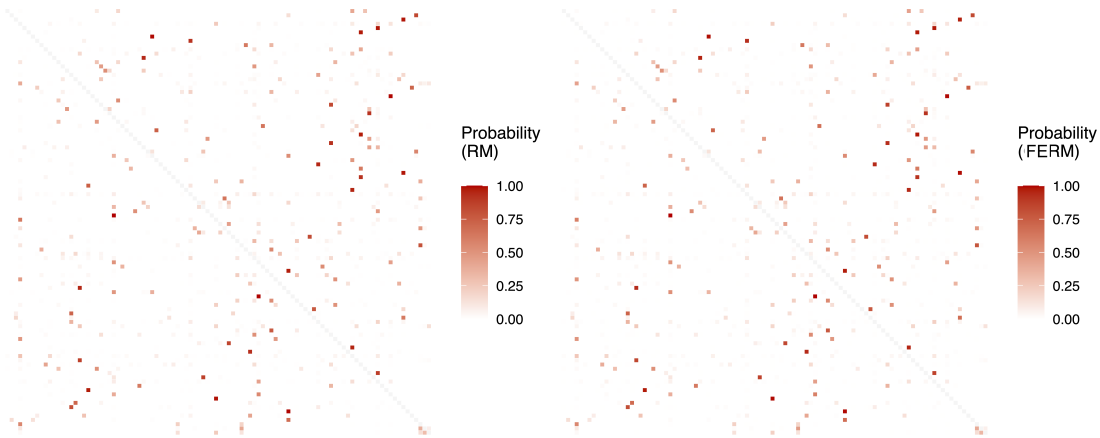
$$J_W\left(A_1, A_2\right) = \begin{cases} \frac{\sum_{i,j \in V} \min\left(a_{ij}^1, a_{ij}^2\right)}{\sum_{i,j \in V} \max\left(a_{ij}^1, a_{ij}^2\right)} & \text{if } \sum_{i,j \in V} \max\left(a_{ij}^1, a_{ij}^2\right) > 0 \\ 1 & \text{if } \sum_{i,j \in V} \max\left(a_{ij}^1, a_{ij}^2\right) = 0 \end{cases}$$

and the Weighted Jaccard distance as

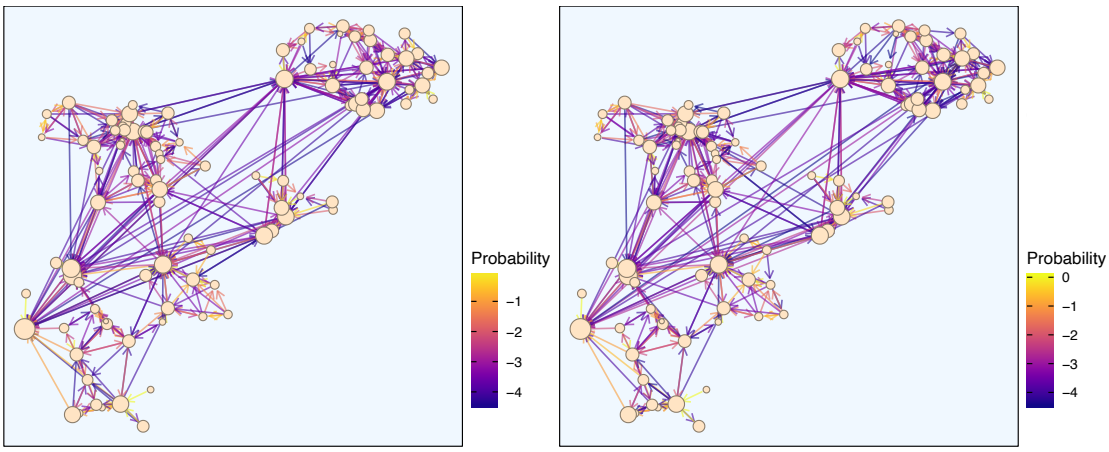
$$d_{WJ}\left(A_1, A_2\right) = 1 - J_W\left(A_1, A_2\right)$$

The Jaccard distance for the two networks is almost zero – the slightly positive value is due to the numerical approximation – indicating the equivalence between the Radiation Model and its generalization.

We illustrate further aspects of the Features-Enriched Radiation Model by taking a plausible spatial distribution of a smaller set of locations as in fig. 6.4. We consider that the benefit for the locations are Gaussian random variables, so that for the origin $p(x_i \mid \theta_i) = p(x_i \mid \mu_i, \sigma_i) = N(x \mid \mu_i, \sigma_i)$ and the same is for the random variable Y, corresponding to the destination. All the nodes have identical benefit distributions except for the nodes $1, 5$ and 8 that are more attractive. In particular, we define $p(x_i \mid 5, 1)$ for $i = \{1, 5, 8\}$ and $p(x_i \mid 0, 1)$ for every other



(a) Probability matrix: each entry i, j corresponds to the probability of a movement from i to j. On the left side: results of the standard Radiation Model (RM). On the right side: results of the features-enriched Radiation Model (FERM). The two matrices are equal.



(b) Networks of flows: the nodes represent the location, and the diameter is proportional to the population. The nodes with higher attractiveness (in the FERM) are highlighted in red. On the left side: the network of the standard Radiation Model. On the right side: the network of the features-enriched Radiation Model. The color of the edges encode the probability of a movement flow between two nodes in a logarithmic scale.

Figure 6.3 – **Results for the synthetic location distribution** of fig. 6.2 with 100 locations

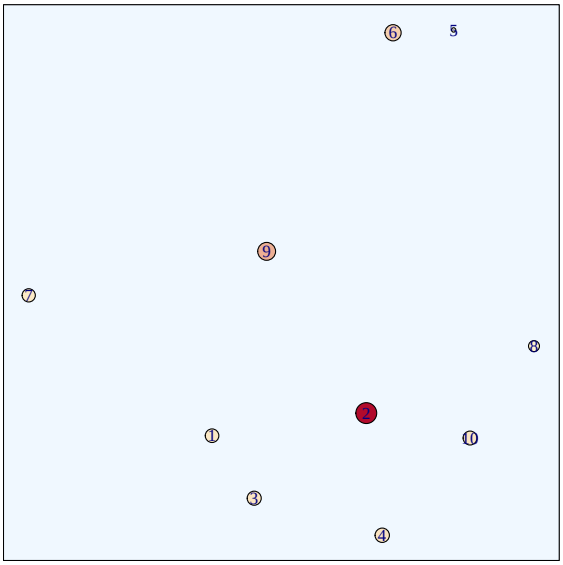
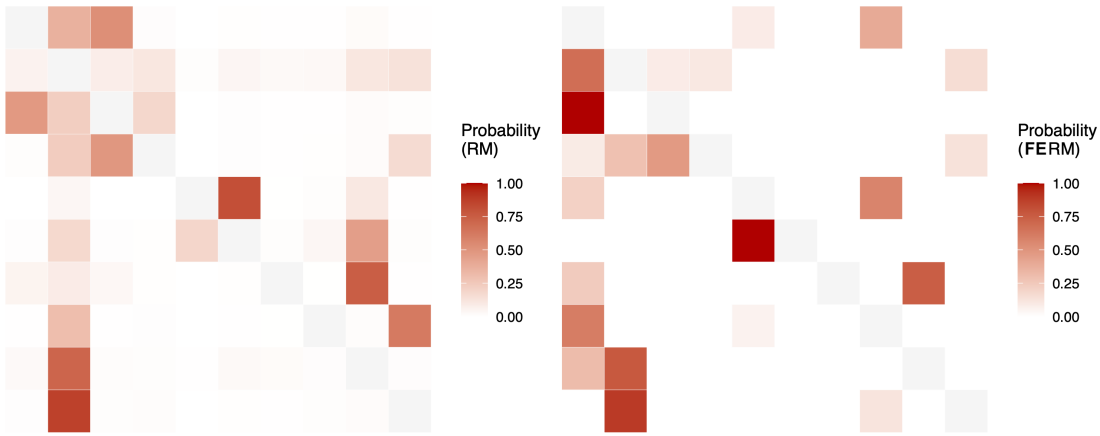
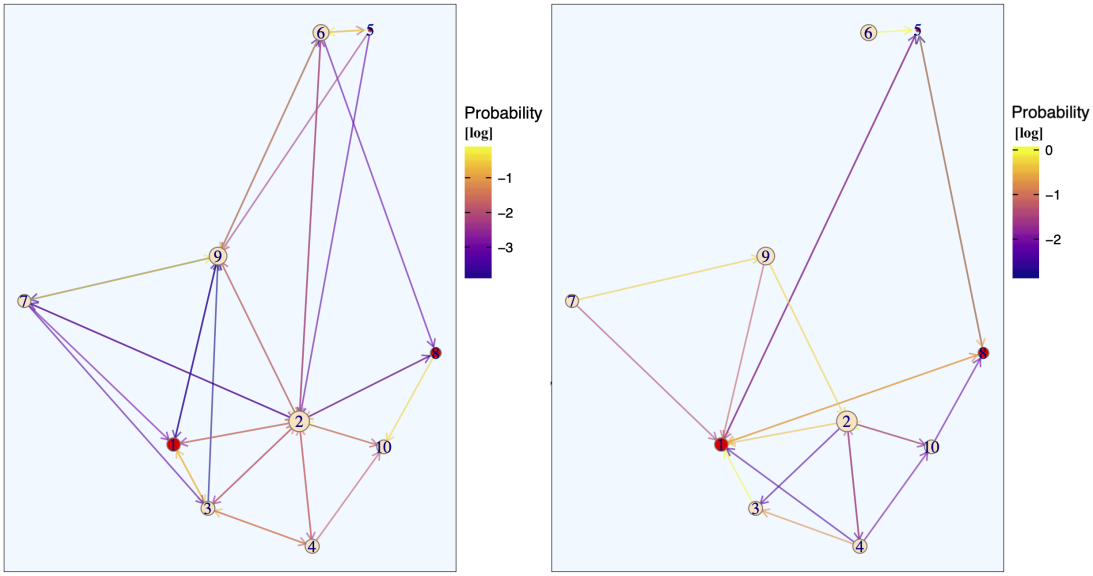


Figure 6.4 – **Synthetic spatial distribution of location** – The coordinates of each location are generated using the Soneira-Peebles model. The color of the circles encode the population at each site.

node. From fig. 6.5 we can draw some considerations about the behaviour of the features-enriched model. First, from fig. 6.5a we can appreciate the redistribution of the probabilities along the columns. In particular, the probabilities on the columns corresponding to the nodes $i = \{1, 5, 8\}$ (with increased attractiveness) are now higher, at the expense of the probabilities in the other nodes (the rows must always sum to 1). Instead, the rows corresponding to the most attractive nodes become whiter since it is not convenient to move away. The same holds true looking at fig. 6.3b, which shows the network structure of the mobility flows. From fig. 6.3b it can be noticed that the flows that were routed towards a node belonging to $\{1, 5, 8\}$ when the attractiveness were homogeneous (RM), now have a sensibly increased probability. For instance, the node 5 is now more attractive, and the flow from node 6 is now directed exclusively there. From node 5 the flows in the RM were oriented towards 2 and 9, in the FERM instead the movements are inhibited since they are no more convenient. A new route is established between 5 and 8 in either directions, since both have increased attractiveness. The FERM inverted the route from 8 to 10, since 8 becomes more attractive although less populated. From 7 the movements stop in 1 rather than going to the more populated 3 or 2. It is to be noticed that, to maintain the continuity also the fluxes between nodes not incurred in increased attractiveness are changed. For example, the route from 9 to 7 disappears. Finally, although the node 2 is the most populated, for the FERM it is no more the most central node (with highest strength), which is instead the



(a) Probability matrix: each entry i, j corresponds to the probability of a movement from i to j. On the left side: results of the standard Radiation Model (RM). On the right side: results of the features-enriched Radiation Model (FERM)



(b) Networks of flows: the nodes represent the location, and the diameter is proportional to the population. The nodes $\{1, 5, 8\}$ with higher attractiveness (acknowledged only by the FERM) are highlighted in red. On the left side: the network of the standard Radiation Model. On the right side: the network of the features-enriched Radiation Model. The color of the edges encode the probability of a movement flow between two nodes in a logarithmic scale. Note the different scale of the color bar.

Figure 6.5 – Results for a synthetic location distribution with 10 locations.

node 1.

These examples illustrate the power of our generalization, that while maintaining the same socio-physical process, it is able to reshape the mobility patters according to the nodes features, making the model flexible and able to adapt to real-world external conditions.

6.4 Scenario Analysis

In this section, we probe the behaviour of the FERM when the features of the geographic locations represent different climatic conditions. Specifically, the features encoded in the model parameters correspond to the climate suitability of different areas under different climate scenarios. The suitability is expressed by changes in the *human climate niche* as defined by Xu et al. in [72]. The socio-economic component of the drivers are still present in the proxy variable *population*, which underpins the computational procedure of the FERM. Furthermore, for the two case studies presented in the following, the GDP – which is commonly used as a socio-economic driver for large-scale mobility – is highly correlated with the population itself, as shown in fig. 6.6, and it would not add further information to the model. The underlying assumption is that the population will change maintaining the same proportion between areas as it is today.

The parameters used in the following experiments are not optimized to fit input data, but rather, are set to prove that the FERM can reproduce mobility patterns that reflect the real-world spatial distribution of location features. We will

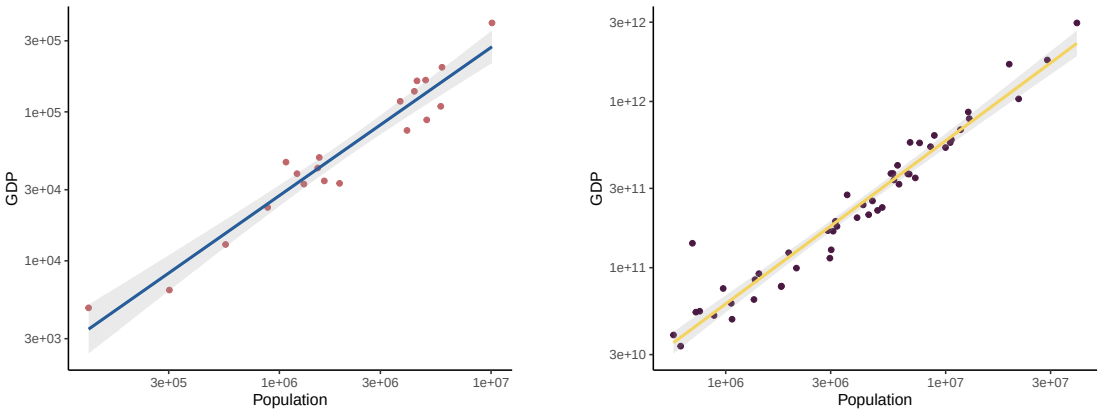


Figure 6.6 – Relation between GDP and Population in the regions of Italy (a) and in the states of the USA (b). These two areas are considered as case studies in the next sections

show that, assuming an optimal parametrization, the Features-Enriched Radiation Model can predict the mobility flows by exploiting exogenous information that a simple RM cannot handle, going beyond the information carried solely by the population. The resulting mobility patterns are consistent with the suitability of the single locations, without neglecting the global effects from the rest of the system.

As discussed in chapters 1 and 5 the environmental, and in particular climatic conditions, may become important drivers of migration. Several climate scenarios are available to quantify the potential changes of the environmental system. In this regard, Xu et al. [72] have taken different climate scenarios to compute the shifts of the human climate niche over time. The climate niche defines the range of environmental conditions in which humans thrive. Specifically, it is defined as the distribution of human population as a function of different climatic variables. The authors have shown that the human climate niche has been surprisingly narrow for thousand of years, since most of the population have concentrated in a limited subset of Earth's available climates, characterized by mean annual temperatures around 13°C. Climate change is expected to alter the position of the niche forcing more than 3 billion people to be left outside the niche, in absence of migration. Here we use the climate niche to drive the FERM and illustrate its potential use for the estimation of the mobility flows under different scenarios. The shift of the climate niche expresses the change in "suitability" of geographic areas, and therefore we use it as a measure of opportunity.

We applied the model to two case studies: Italian regions and USA countries. In both cases, the climatic conditions varies widely over the area, making the results particularly informative. For both the case studies, we focused on two climate niche scenarios hereinafter called *Temperature* and *Temperature + Precipitation* scenarios, that have been derived under the business-as-usual scenario for climate (Representative Concentration Pathway 8.5, RCP8.5) and the "high challenges" socioeconomic scenario (Shared Socioeconomic Pathway 3 [SSP3]) for the population growth in the absence of significant migration [72]. The first scenario considers only the temperature to define the climate niche, whereas the second one takes into account both temperature and precipitation (we refer to [72] for more details on the scenarios generation).

6.4.1 Inter-regional mobility in Italy

Italy stretches across the centre of the Mediterranean, from a latitude of 36°N to a latitude of 47°N. This remarkable extension, traversed by a peculiar orography, leads to iridescent climatic conditions. The main climatic regions are: (i) the Alpine Region, (ii) the Po Plain and Upper Adriatic Region, (iii) the Central-Southern Adriatic Region, (iv) the Ligurian-Tyrrhenian Region, (v) the Apennine

Region and (vi) the Mediterranean Region [363]. In the widely adopted Köppen classification, Italy falls within the Mediterranean climate area, which is part of the subtropical climates with dry summers. Climate change is expected to alter these conditions, and consequently, the suitability of the regions will change accordingly. The change in suitability due to the shift in the climate niche is represented in fig. 6.7. This figure shows that Italy will be divided into two parts:

- the territories of the northern Italy and in particular the alpine regions will become more suitable, since at higher temperature the mountains of Alps will be warmer;
- the southern Italy will instead face an environmental deterioration due to higher temperature and water scarcity [364, 365].

The conditions under the *Temperature + Precipitation* scenario are noticeably

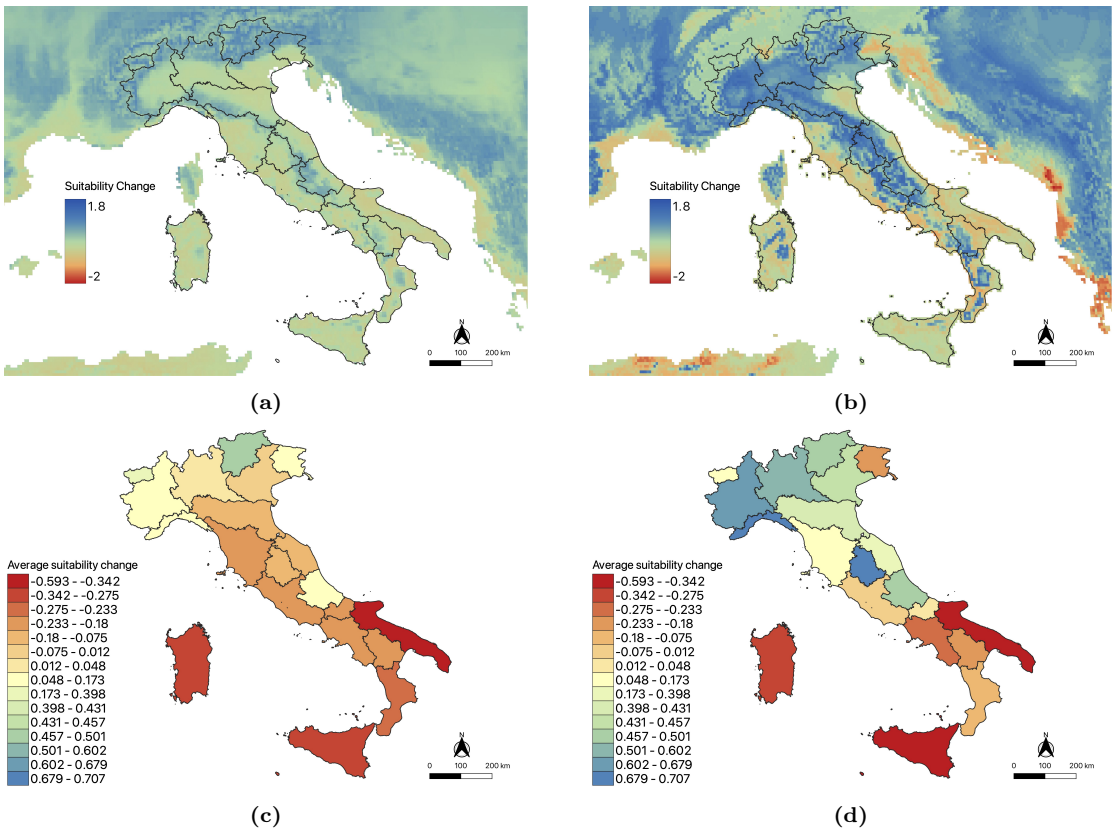
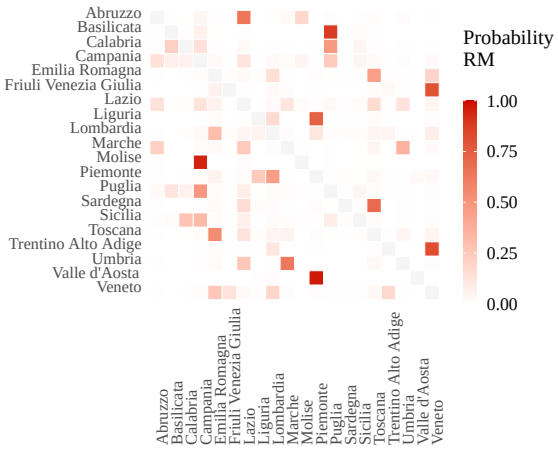


Figure 6.7 – Climate Niche Difference – Italy – (a) Scenario: *Temperature*; **(b)** Scenario: *Temperature + Precipitation*; **(c)** Average niche difference under Scenario: *Temperature*; **(d)** Average niche difference under Scenario: *Temperature + Precipitation*. The climate niche difference is computed between current and the future climate scenario RCP8.5

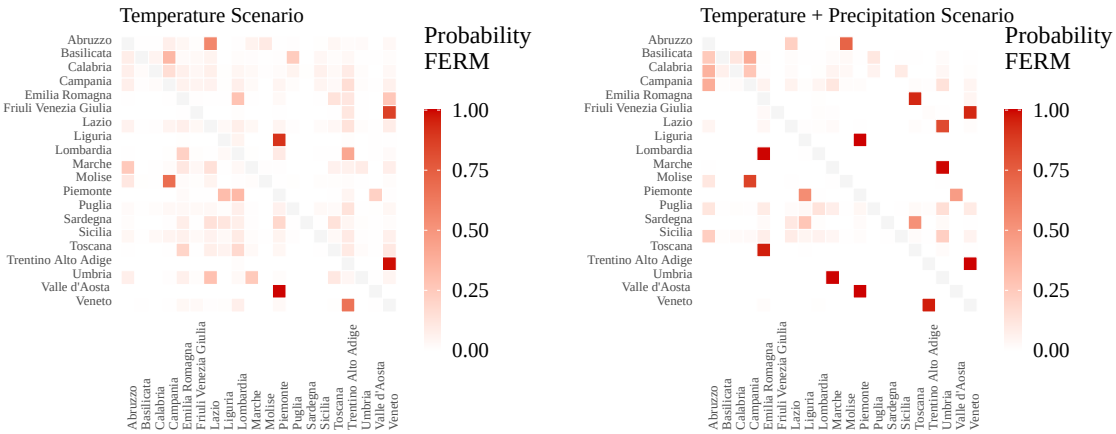
Figure 6.8 – **Average niche shifts – Italy** – The values reported in this figure are also used as the mean parameter for the benefit distributions of each location.

more extreme in both positive and negative directions, but are overall more favourable than the *Temperature* scenario (the average niche shift for *Temperature + Precipitation* in Italy is around +0.15). In both the scenarios some regions become more suitable, whereas other ones encounter worse conditions. We use the average niche shifts to drive the FERM over the Italian regions. The average values are reported in fig. 6.8. In particular, the mean of the benefit distributions are fixed equal to the average niche shifts of the related location, while the variance is set to 1 in every location (see *Appendix* for the results of the sensitivity analysis). The results of the model are shown in the OD matrices in fig. 6.9. These matrices give the probability of migration flows from the regions on the rows to the ones on the columns. This representation highlights the ability of the FERM to redirect the flows changing the migration patterns. It can be noticed that some mobility routes remain open, but with a different probability, while other new ones appear. In fig. 6.9b) we see that the migration pattern changes drastically, as a consequence of the more extreme exogenous conditions of the *Temperature + Precipitation* scenario.

If the modification of the mobility patterns are clearly illustrated by the OD matrices, the fig. 6.10 gives a more accurate insight on the causes of this modification. On the top of the figure is reported the barplot already seen in fig. 6.8 for



(a)



(b)

Figure 6.9 – Probability OD matrices – Italy – Each entry (i, j) correspond to the probability of a flow between the locations i and j, as computed by RM **(a)** and FERM **(b)**. Although the FERM is able to redirect the flows changing the migration patterns, some routes remain open, but with a different probability. On right hand side of **(b)** we see that the migration pattern changes drastically, as a consequence of more extreme exogenous conditions under *Temperature + Precipitation* scenario.

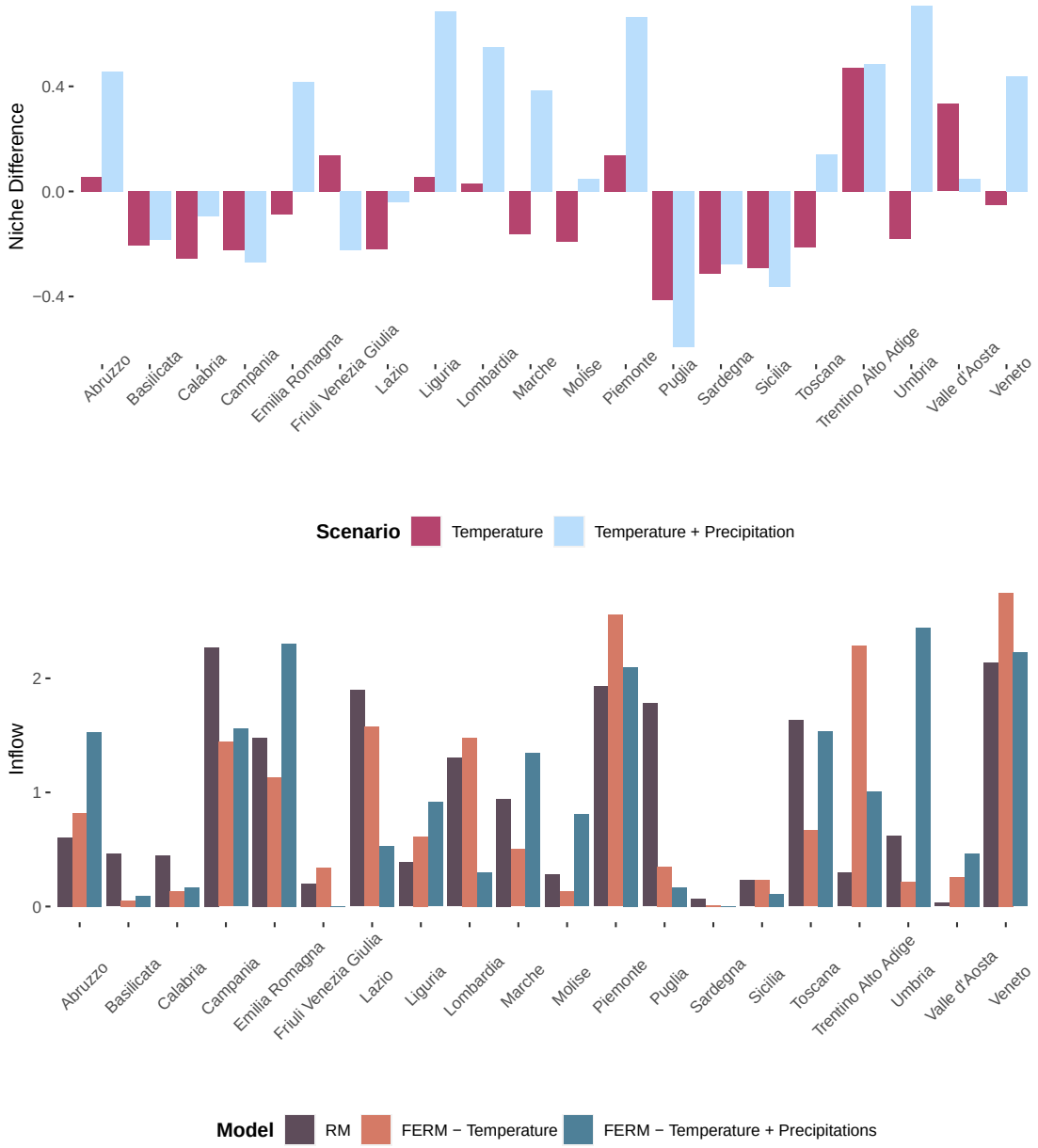


Figure 6.10 – Scenario analysis aggregated results for Italian regions

easier comparison with the flows estimated under the two scenarios by the FERM and the RM. The bars plotted in the figure below refers to the aggregated probability inflows for each region. In general, the inflows are expected to grow if the region will face better conditions, and vice versa. The model works as expected, with some interesting additional properties. Specifically, the figure shows that regions such as Abruzzo, Basilicata, Calabria, Campania will attract less fluxes in response to worsening conditions, and vice versa depending on the scenario. Also Emilia-Romagna gives an evident example of this behaviour, where the two scenarios indicate opposite conditions the mobility flows follow the same pattern. The same happens for Trentino-Alto Adige and Marche. In both the scenarios, Puglia turns out to be the region facing the worst conditions, indeed the flows are greatly reduced following the same worsening pattern as the climate niche.

Some regions return interesting and unexpected results. In Piemonte, for instance, the climate niche is likely to improve in both scenarios (this region is vastly occupied by mountains of the Alps) where under the *Temperature + Precipitation* scenario the improvement is much larger. The fluxes instead, are smaller for this latter scenario, although increased with respect to the RM. In Sicily, instead, the climatic conditions are expected to heavily worsen, but the fluxes will remain quite stable. Finally, the case of Lombardia is peculiar: in the *Temperature* scenario the suitability will slightly increase and the inflows will be larger as well, but under the great increase of suitability under *Temperature + Precipitation* scenario the inflow instead becomes much smaller than both RM and the *Temperature* scenario.

These unexpected behaviours suggest that the model is driven not only by the direct influence of the location features, but also by the global conditions of the system. For the specific case of Lombardia, it should be noticed that this region is surrounded by regions of similar suitability, especially in the *Temperature + Precipitation* scenario (see fig. 6.7), so that the inflow may be reduced by the fact that from the nearest regions the outflow are inhibited by the higher suitability, albeit Lombardia itself become more attractive.

The consideration exposed so far, can be quantified by computing the correlation between the average niche change and the difference between the flows from RM and FERM. In this way one can measure the effect of the niche change on the mobility patterns. The results of this procedure are graphically represented in fig. 6.11. In this figure it is clear the association between the variables, and the unexpected cases described above appear as outliers. The statistical relation is further corroborated by testing the estimated Spearman's rank correlation reported in table 6.1. The correlation is computed also on the sign of the variables to obtain an indicator of how often increased suitability corresponds to increased inflow. All the correlation are statistically significant, and the estimated values confirm that the model can interpret the meaning of the features through the

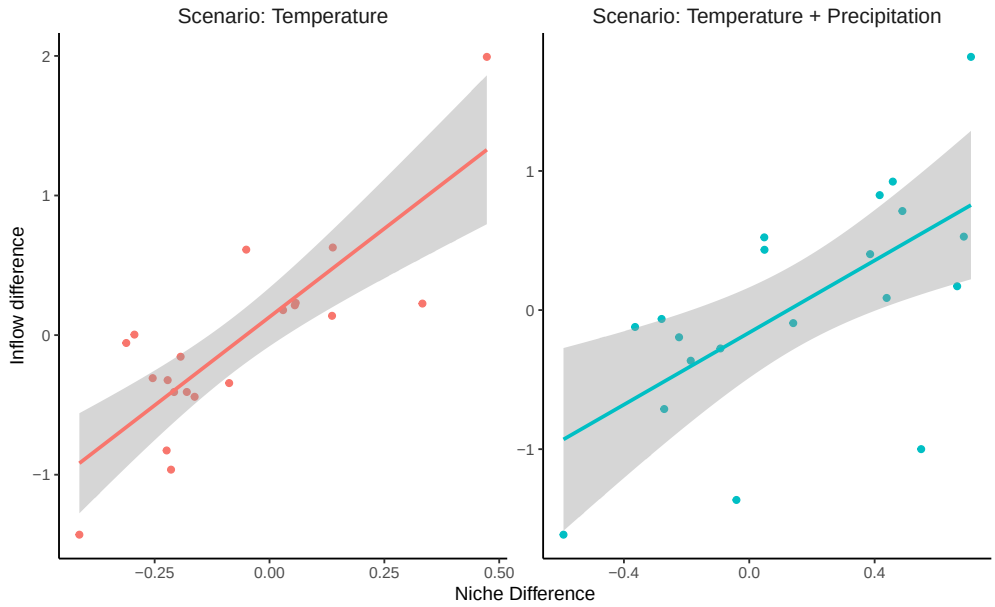


Figure 6.11 – Relation between climate suitability and FERM results in the two scenarios – Italy. On the x-axis is the niche shift between current and future climate; the y-axis represent the difference between the inflow computed with the FERM and the RM. The solid lines are the linear regressions.

	Spearman Corr.		Spearman Corr. on signs	
	estimate	p-value	estimate	p-value
Niche shift - Inflow Difference (*Temp*)	0.720	0.00050	0.811	1e-05
Niche shift - Inflow Difference (*Temp + Prec*)	0.642	0.00286	0.816	1e-05

Table 6.1 – Results of the correlation tests between niche shift and the difference in inflow between FERM and RM.

parameters, and govern the flows accordingly. Thus, the model is able to direct the mobility flows towards the most attractive nodes, balancing the information carried by the population.

The OD matrices reported in fig. 6.9 can be represented with the mobility networks of fig. 6.12. This latter figure summarizes the results of the FERM driven by the climate niche. The differences between the RM and the FERM can be appreciated, together with the features of the regions. The network generated by the FERM under the *Temperature* scenario (fig. 6.12(b)) is more dense than the one from the RM (fig. 6.12(a)), but the fluxes have in general larger magnitude. This is due to the fact that the climatic conditions of southern Italy get worsen, and the fluxes form north to south have disappeared. The network for the scenario *Temperature + Precipitation* is even less dense, because the fluxes are inhibited

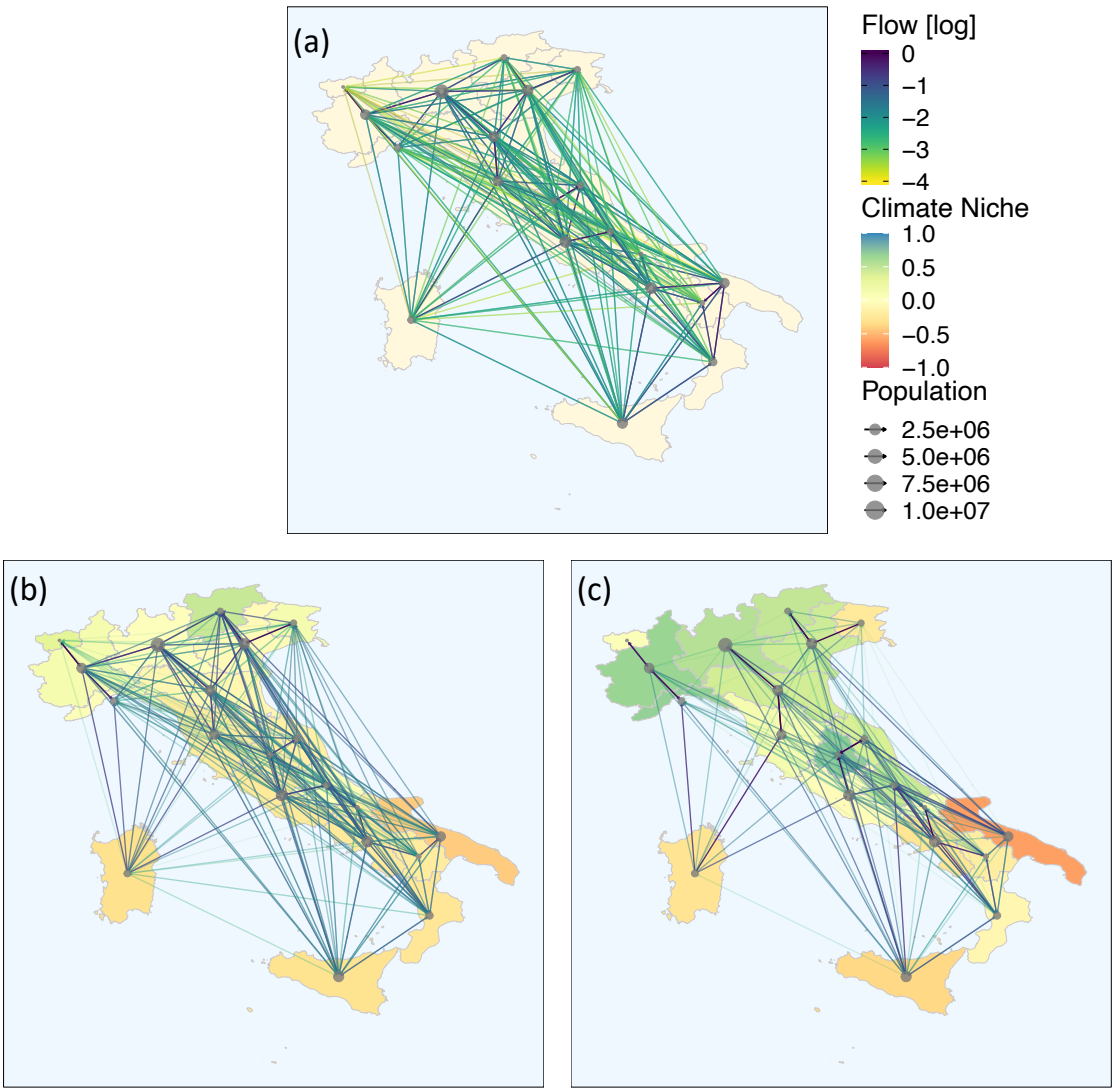


Figure 6.12 – **Mobility Networks** – **Italy** – as computed by RM **(a)** and FERM under the scenarios *Temperature* **(c)** and *Temperature + Precipitation* **(c)**.

by a generally better climatic conditions (the climate niche shift is around +0.15), albeit the fluxes from the southern to the northern regions are still present and are intensified.

6.4.2 Inter-state mobility in USA

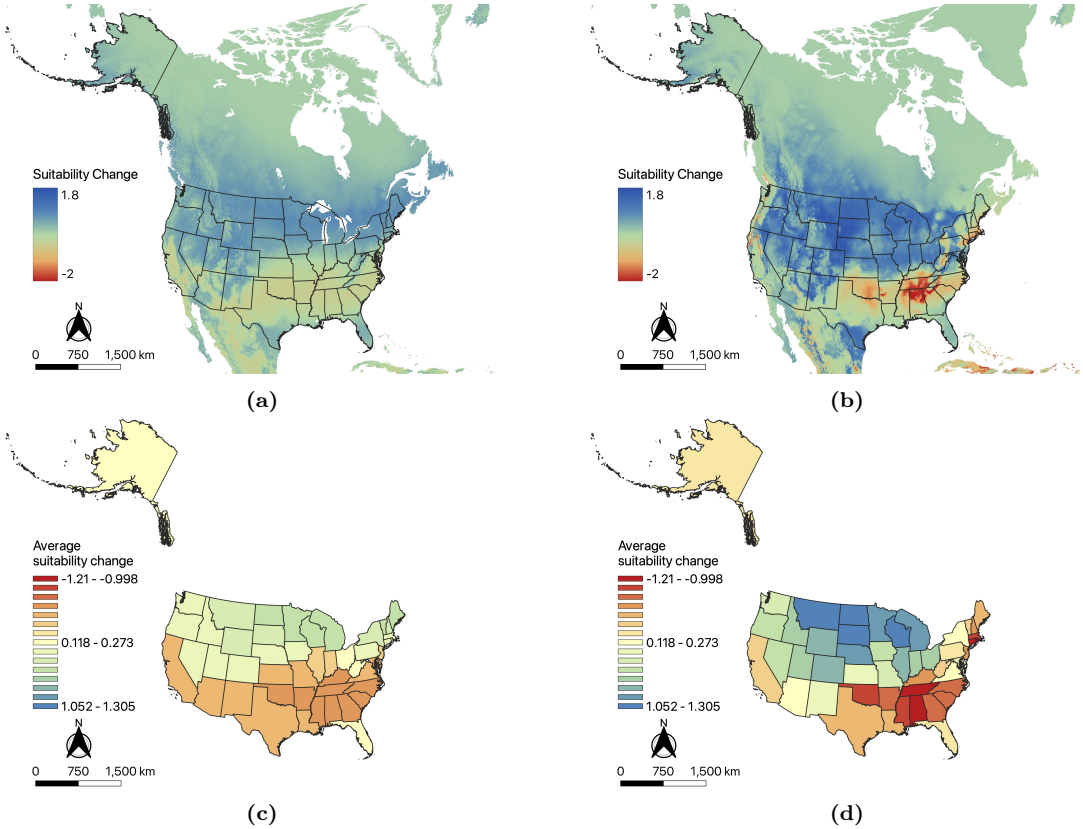


Figure 6.13 – Climate Niche Difference – USA – (a) Scenario: *Temperature*; **(b)** Scenario: *Temperature + Precipitation*; **(c)** Average niche difference under Scenario: *Temperature*; **(d)** Average niche difference under Scenario: *Temperature + Precipitation*. The climate niche difference is computed between current and the future climate scenario RCP8.5 and assuming a the population density as given by the socio-economic scenario SSP3.

The same analysis has been conducted also for the case of the United States of America. This area encompasses a wide variety of climatic conditions, spanning from mountains to deserts. Generally, the climate of the U.S. is warmer in the south, drier in the west, and alpine in some regions of the north and west. These characteristics, although will endure, are expected to be exaggerated and relocated by climate change. The climate niche shift in 2070, as reported by the two considered scenarios, is represented in fig. 6.13 As happened for the Italian

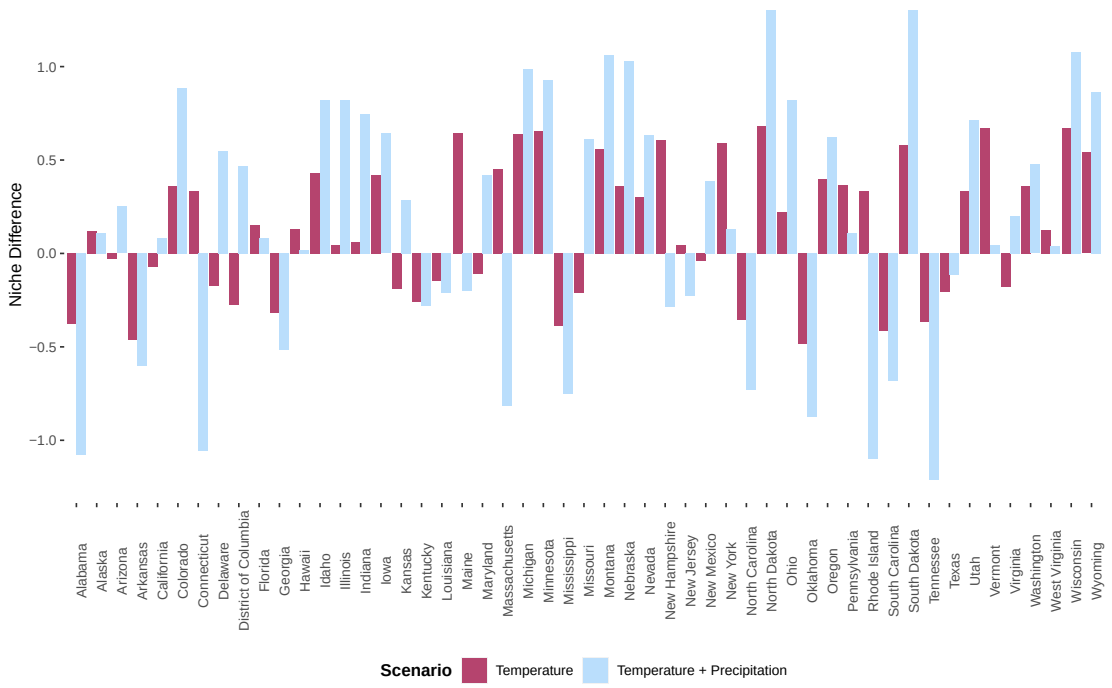


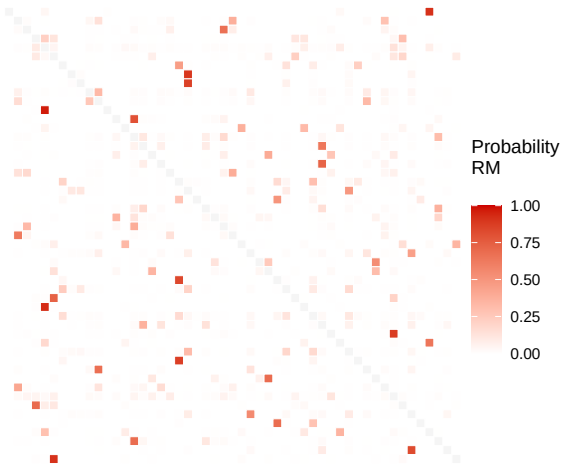
Figure 6.14 – Average niche shifts – USA – Each bar indicates the value of the climate niche shift averaged over a state. The values reported in this figure correspond to the mean parameter for the benefit distributions of each location.

regions, also for the USA the *Temperature + Precipitation* scenario gives the most extreme conditions, on both positive and negative sides of the scale.

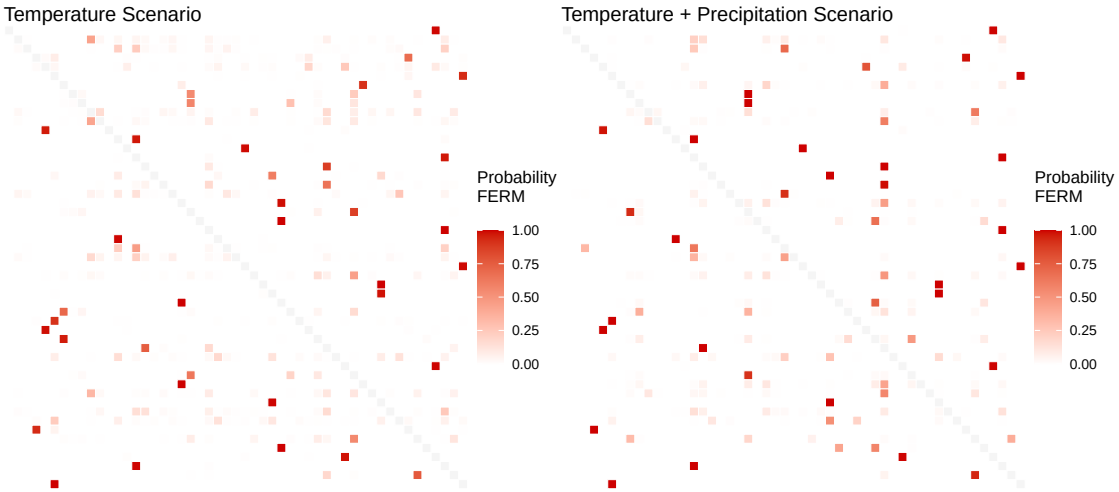
The average climate niche shift over the states are used again as the feature to drive the FERM. The values, for the two scenarios are reported in fig. 6.14. The countrywide conditions are slightly enhanced in the two scenarios with an average niche shift of +0.14 for the *Temperature* scenario and +0.17 for the *Temperature + Precipitation* scenario. Nonetheless, there are states – such as Alabama, Tennessee and Oklahoma – in which the conditions get much worse.

In fig. 6.15 are reported the OD matrices representing the probability of the mobility flows between the states. Although the spatial scale is way larger than the previous case study, the results are qualitatively similar to the ones obtained for the Italian regions. In fact, the overall shift in the climate niche for the two areas are similar. In both the cases there are clusters of areas with similar changes in suitability that exchange mobility fluxes with each other.

The results in term of aggregated inflows are reported in fig. 6.16. Also in this case, the effect of a shift in the climate niche produces both local foreseeable effects and unexpected behaviours. The inspection can be made by analysing the fig. 6.16



(a)



(b)

Figure 6.15 – Probability OD matrices – USA – Each entry (i, j) correspond to the probability of a flow between the locations i and j, as computed by RM **(a)** and FERM **(b)**. As for the previous case study, the FERM is able to redirect the flows changing the migration patterns, still maintaining some routes open, but with a different probability. Also in this case, the scenario that consider both temperature and precipitation is more extreme, leading to drastic changes in human mobility as shown on the right plot of **(b)**.

or by relying on fig. 6.17 which shows – as before – the relation between niche shift and the difference of flows between FERM and RM. Most of the aggregated flows increase with improving climate conditions, giving an high value of correlation (see also table 6.2), confirming even at larger scale that the model can control the mobility fluxes directing them towards the most attractive locations, balancing the effects of the population concentrated in the nodes.

The favourable conditions of the South Dakota, which is also in the middle of the country, leads to one of the highest relative increase in the inflows in both the scenarios. Under the *Temperature + Precipitation* scenario, the Tennessee is the most afflicted state, and this is reflected into the almost complete suppression of the inflows. One of the outliers can be recognised in the state of Ohio, that in response to a not very high increase in the niche, faces the highest increase in the inflow, under the *Temperature + Precipitation* scenario. This is arguably due to its proximity with highly populated nodes that will encounter instead a decline in the climatic conditions.

The results are again summarized by the mobility networks in fig. 6.18, in which are shown also the climate niche changes. As in the previous case study, under the *Temperature* scenario the network is denser, ad mostly similar to the one given by the RM. The conditions are in fact quite homogeneous and the influence of the population is higher. The short-range flows remain crowded while the long-range are dampened by the general rise in the climatic conditions. The *Temperature + Precipitation* scenario leads again to a "sparsification" of the network, since on average the conditions are enhanced (the average niche shift is +0.17). The long-range flows are those towards the highest climate suitability, despite the low population – e.g. from Alabama to North Dakota. This figure puts in evidence the appearance of massive flows between the states of the New England and Ohio, as discussed before. In general, the states in adverse conditions on the south-east increase their outflows towards the north-west.

In conclusion, we showed the potential use of the FERM in predicting mobility patterns using the features of the nodes as drivers of migration. Nevertheless, optimized parameters and further exogenous information should be used to get plausible scenario more adherent to the real world. We should note that, the climate niche takes into account just two global variables, and neglect important climate factors such as the occurrence of extreme events due to climate change, or slow onset environmental changes such as the sea level rise. Catastrophic environmental events, as already discussed in chapters 1 and 5 may become an increasingly significant driver of migration. For this reason, further studies should apply the FERM by including (at least) these additional features in the model. Other variables from the socio-economy considered as causes of migration, such as social connectedness, cultural identity/diversity and other historical features may be similarly included.

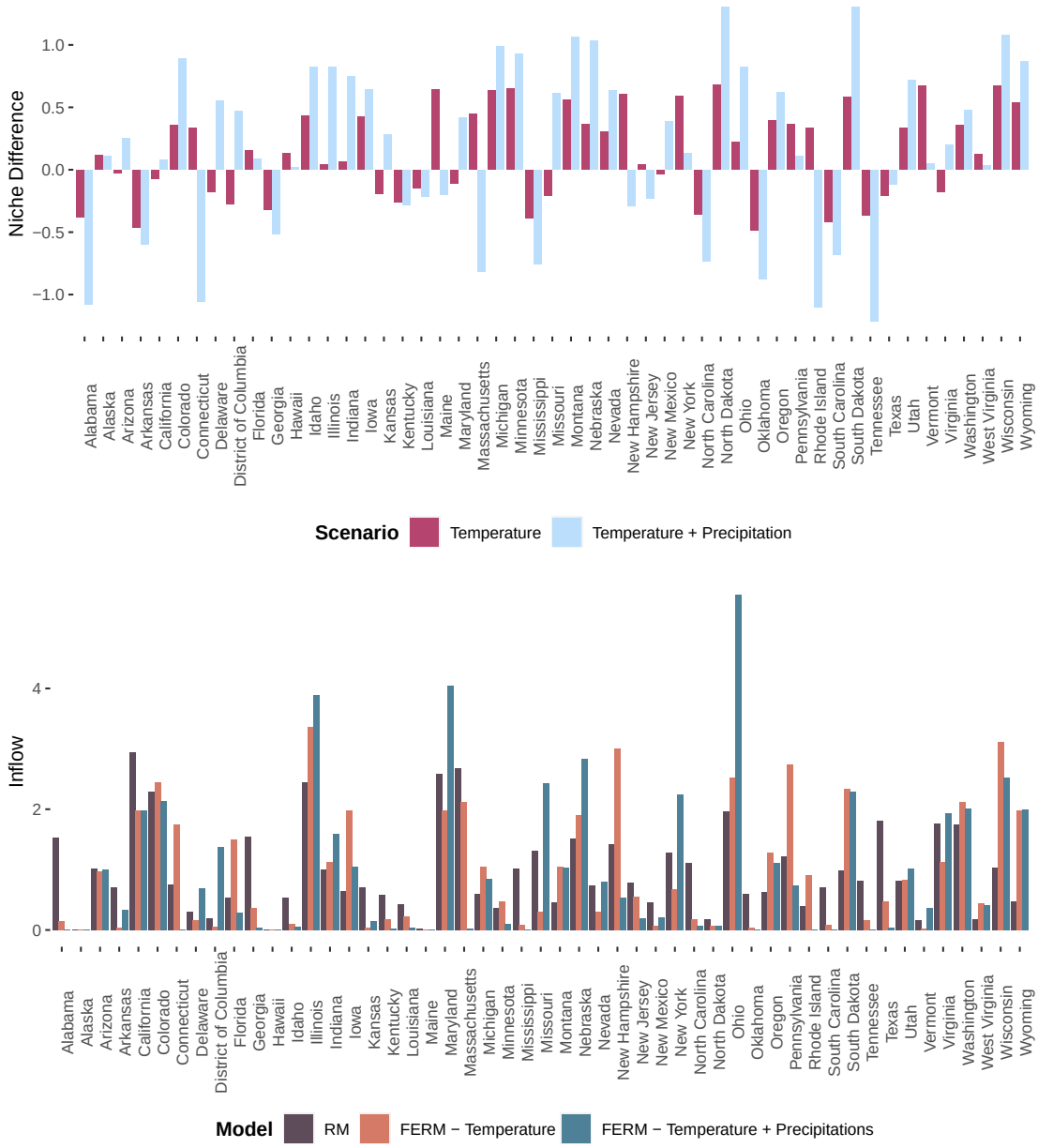


Figure 6.16 – Scenario analysis aggregated results – USA

	Spearman Corr.		Spearman Corr. on signs	
	estimate	p-value	estimate	p-value
Niche shift - Inflow Difference (Temp)	0.684	1.1e-07	0.672	6.685e-08
Niche shift - Inflow Difference (Temp + Prec)	0.746	$<$ 2.2e-16	0.616	1.505e-06

Table 6.2 – Results of the correlation tests between niche shift and the difference in inflow between FERM and RM.

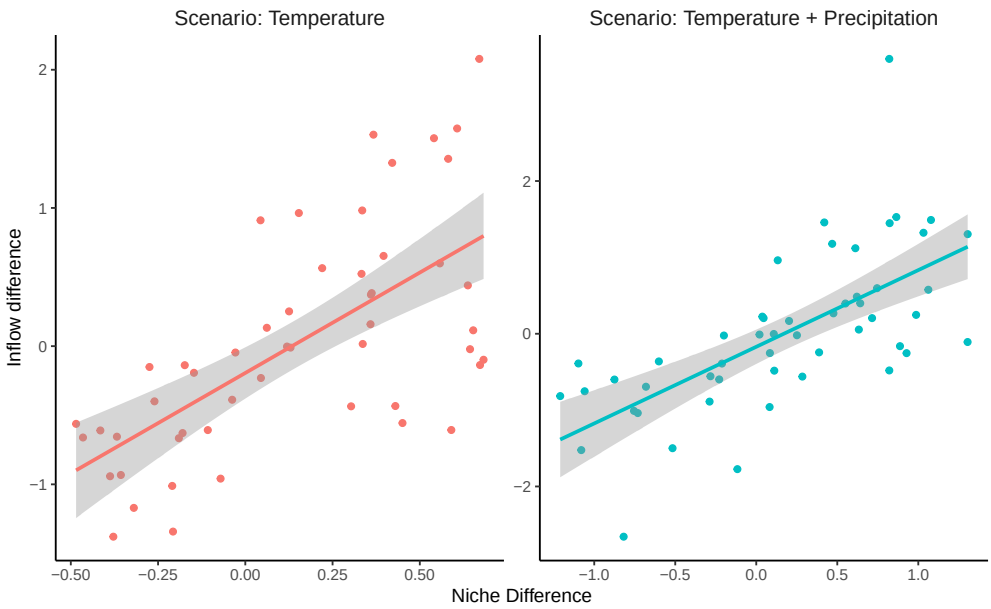


Figure 6.17 – Relation between climate suitability and FERM results in the two scenarios – USA. On the x-axis is the niche shift between current and future climate; the y-axis represent the difference between the inflow computed with the FERM and the RM. The solid lines are the linear regressions.

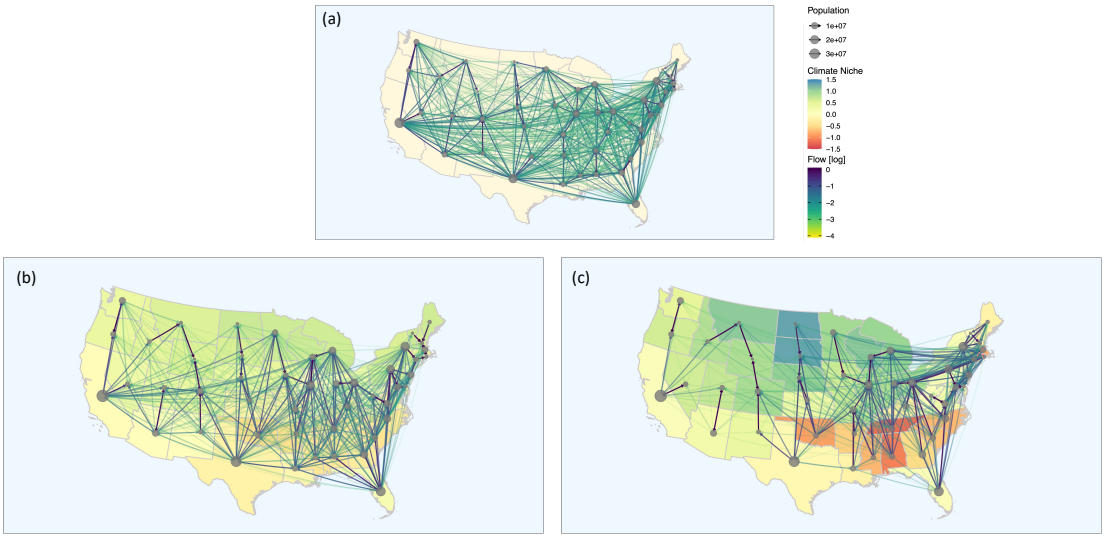


Figure 6.18 – Mobility Networks – USA – as computed by RM **(a)** and FERM under the scenarios *Temperature* **(c)** and *Temperature + Precipitation* **(c)**. The edges linking Alaska and Hawaii with the continent have been left out for display purposes.

6.5 Conclusions and future research directions

In this chapter we presented the Features-Enriched Radiation Model, a mathematical generalization of the Radiation Model for human mobility. The Features-Enriched Radiation Model expands the spectrum of possible application of the Radiation Model to those cases in which, besides the population, other exogenous information about the attractiveness of the locations are modelled as *causes* of large-scale human movements. Our model maintains the same physical process of the original model, but generalizes the main mathematical constituents to accommodate the locations features. The features act on the probability of establish a mobility flows between locations, by modifying the parameters of the *benefit distributions* of the Radiation Model. These distributions are thus tailored to each location making them more or less attractive, balancing the effects of populations, that are still present as in the original model. We demonstrated the equivalence between our model and the Radiation Model in the case in which the locations are considered indistinguishable. To accomplish this task we built a "synthetic world" with a realistic spatial distribution of location, and we showed the main characteristic of the model on a smaller scale. Subsequently, we fed the model with real-world features derived from the *human climate niche*, for two areas: Italy and United States of America. With these two case studies we proved that the FERM can reproduce mobility patterns that reflect the spatial distribution of location features.

Further studies are needed to make the FERM suitable for predicting human mobility flows in a broad range of situations. First of all, it requires an efficient routine for the calibration on real data. To this purpose, a promising approach would be the use of a "Bayesian machine scientist", as defined in [366], that could be able to estimate not only the model parameters, but also the most appropriate form of the benefit distribution. To calibrate the model, historical data of bilateral mobility are needed. This is currently a general limiting factor for human mobility modeling. Having enough data, an interesting application of the FERM is on the spatial scales and geographical areas in which the original RM is known to fail (for example at the scale of cities and in low-income countries [357]). The model can be used also to discern the different causes of migration, by making an appropriate model selections.

Appendix

"Synthetic worlds" generation

To test the model against the early Radiation Model, we generated a set of "synthetic worlds", namely distributions of points in a two-dimensional space, on the wake of [367]. To this aim, we implemented an algorithm that provides an autosimilar spatial distribution of points with fractal dimension. The algorithm, conceived by Soneira and Peebles in 1978 [361], was originally used to describe the spatial distribution of galaxies in the universe. It provides a spatial distribution to a given number of points. The algorithm proceeds by successive steps, placing a set of spheres or circles in space. The level-0 circle is of radius R. Inside the level-0 circle a number η of level-1 circles with radius R/λ with $\lambda > 1$ are placed at random. Analogously, within the level-1 circles, another η level-2 circles of radius R/λ^2 are randomly placed. This process is repeated until a number of ηL circles of radius R/λ^L are arranged. At the end of the process, at the center of each circle is placed a point, that will correspond to the geographical location used in the FERM.

The model depends on the three parameters λ, η and L, that together con-

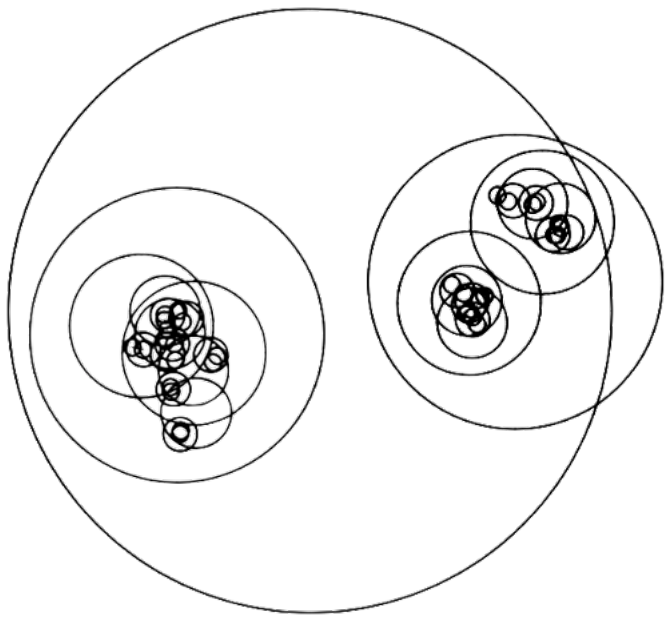


Figure 6.19 – **The Soneira-Peebles model** – Inside a level-0 sphere η level-1 circles are placed with a radius which is smaller by a fixed factor. This process is repeated until one ends up with ηL level-L circles. At the center of these level-L circles ηL points are placed, which form the resulting Soneira-Peebles point distribution. Figure and caption from [368].

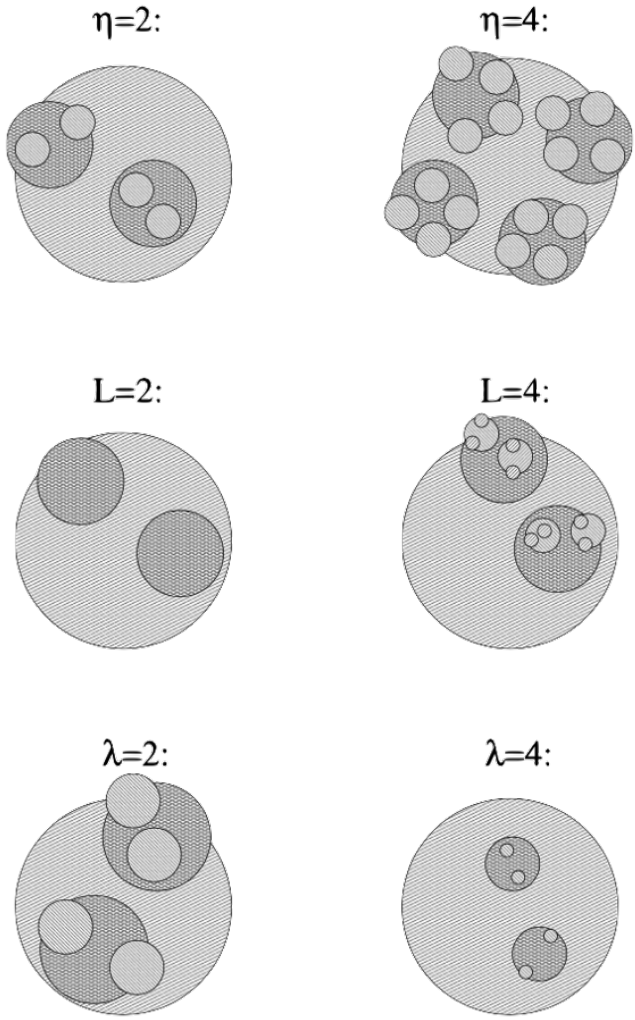


Figure 6.20 – The physical meaning of the three defining parameters η, L and λ of the Soneira-Peebles model – The upper row shows the effect of varying η, the number of circles which is placed in each circle. The central row shows the effect of varying L, the total number of levels. The bottom row shows the effect of varying λ, the ratio of the radius of each circle with the radius of subsequent circles of on. Figure and caption from [368].

tribute in determining the fraction of space covered by the circles, the size of the circles, the density and scale in the resulting distribution and its fractal dimension (see figs. 6.19 and 6.20). Importantly, the two-point autocorrelation function for the spatial distribution is given by

$$\xi(r) \sim r^{-\gamma} \quad \text{with} \quad \gamma = M - \left(\frac{\log \eta}{\log \lambda}\right) \quad \text{for} \quad \frac{R}{\lambda^{L-1}} < r < R$$

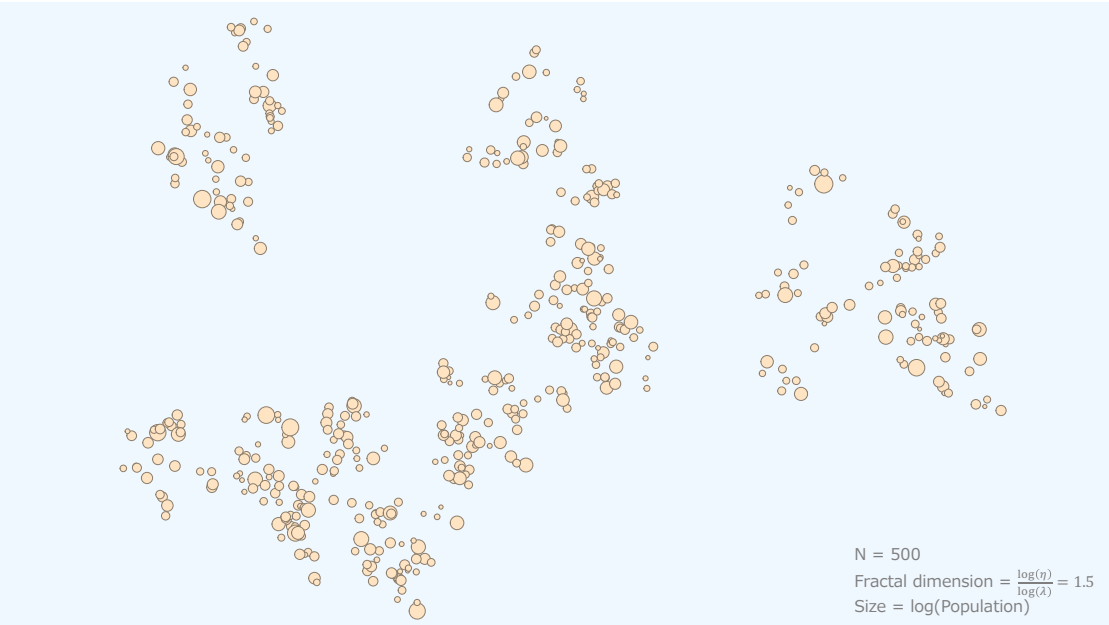


Figure 6.21 – Example of a synthetic world generated with the Soneira-Peebles model

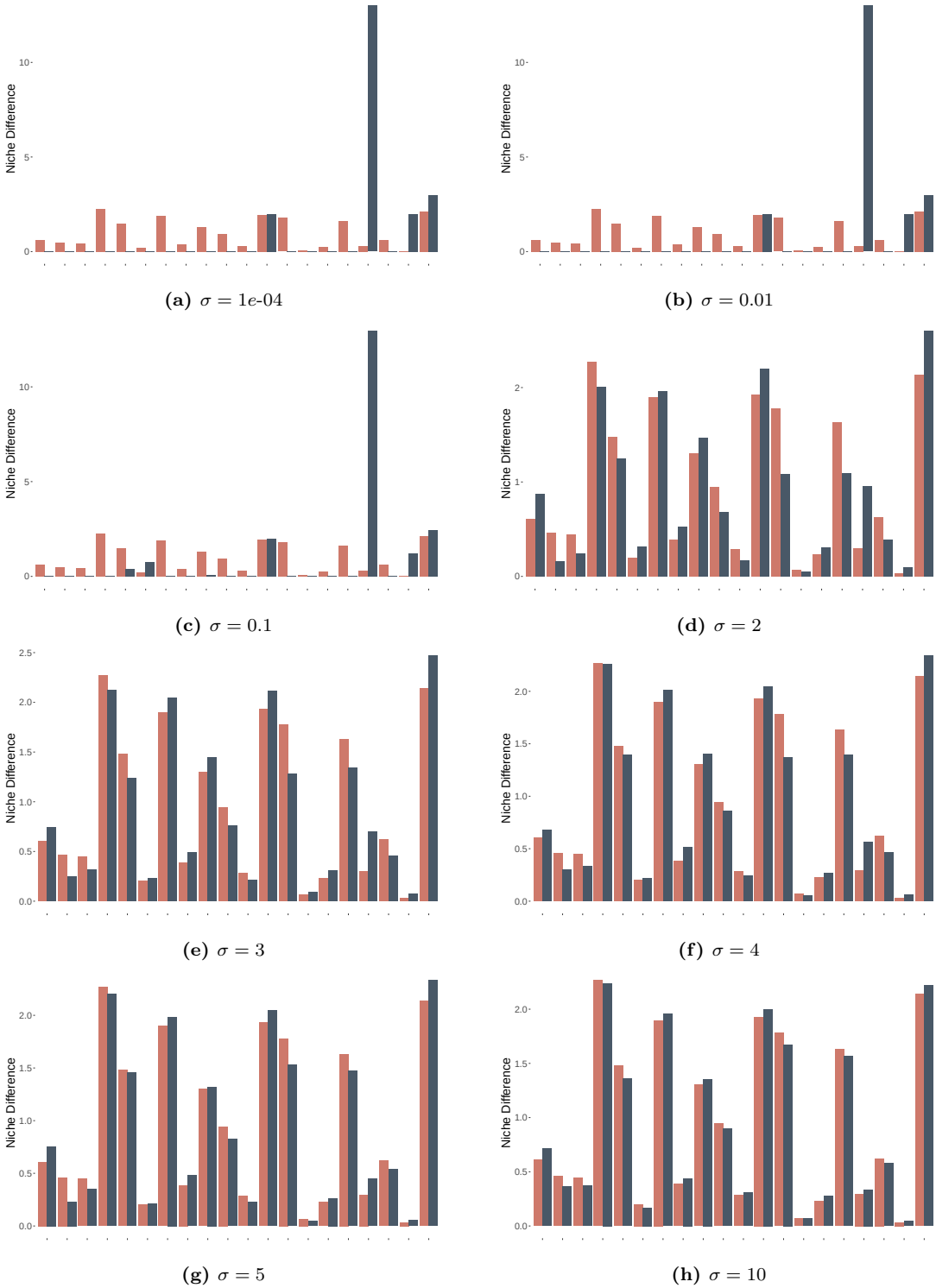
and the fractal dimension is defined as

$$D = \lim_{r \to 0} \frac{\log N(r)}{\log(1/r)}.$$

Besides the examples in the main text, an illustration of a spatial distribution generated with the Soneira-Peebles model is depicted in fig. 6.21.

Sensitivity analysis on variance

For the two case studies of section 6.4 we fixed the variance of the benefit distribution to 1. This value is a good balance between two extremes: when the variance tends to zero, the radius of exploration around a location tends to zero. Therefore, the fluxes are kept only between locations with overlapping benefit distributions, and geographically close to each other. On the contrary, when the variance tends to infinity, all the benefit distributions are flat, and they converge to a unique uniform distribution, restoring the same conditions of the simple Radiation Model, in which the benefit distributions are assumed to be equal for all the locations. This is illustrated in the figs. 6.22 and 6.23, for the case study of the Italian regions and for both the scenarios.



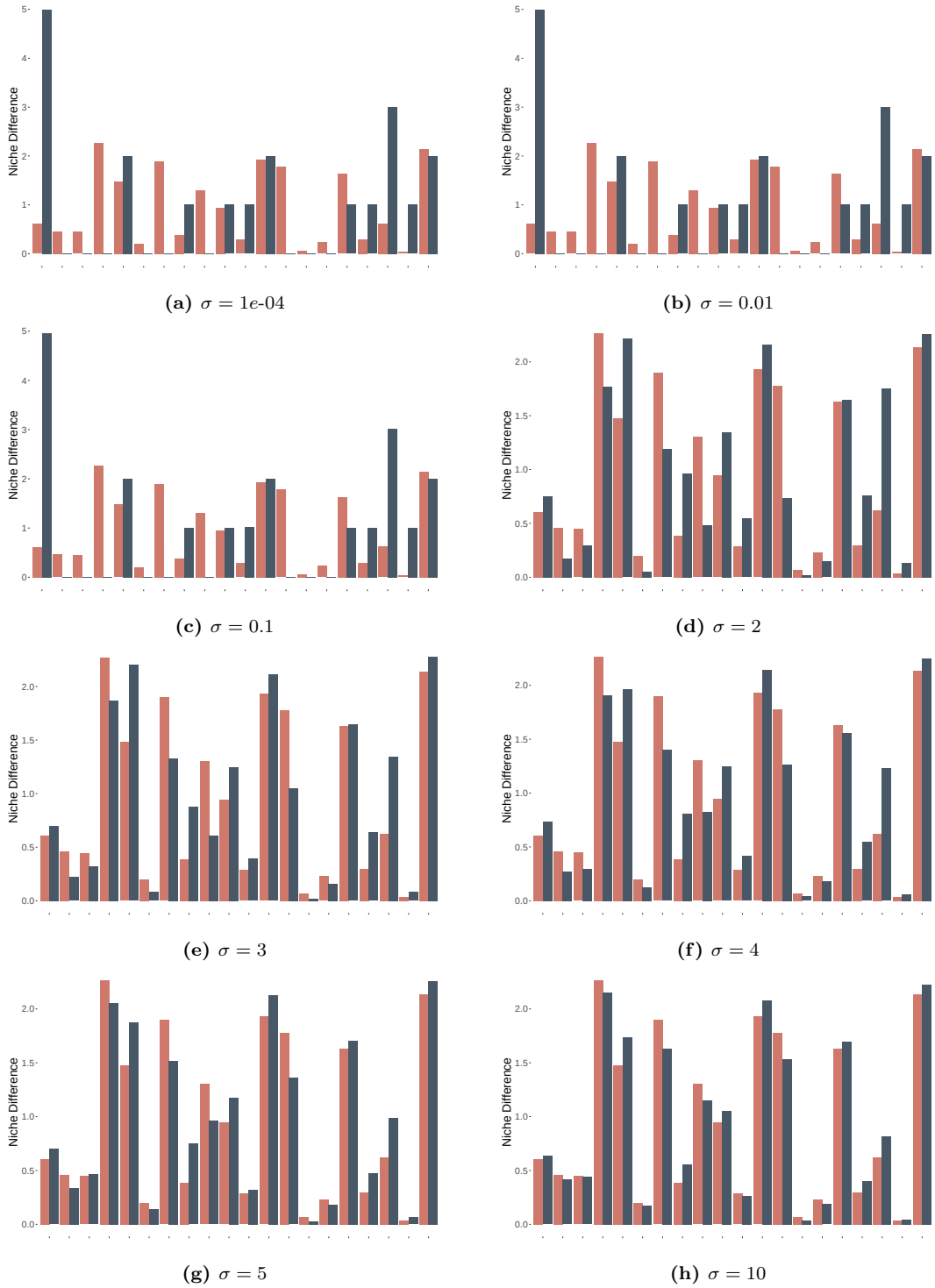
(a) $\sigma = 1e\text{-}04$ **(b)** $\sigma = 0.01$

(c) $\sigma = 0.1$ **(d)** $\sigma = 2$

(e) $\sigma = 3$ **(f)** $\sigma = 4$

(g) $\sigma = 5$ **(h)** $\sigma = 10$

Figure 6.22 – Sensitivity analysis for Italian Regions – Scenario *Temperature*



(a) $\sigma = 1e\text{-}04$ **(b)** $\sigma = 0.01$

(c) $\sigma = 0.1$ **(d)** $\sigma = 2$

(e) $\sigma = 3$ **(f)** $\sigma = 4$

(g) $\sigma = 5$ **(h)** $\sigma = 10$

Figure 6.23 – Sensitivity analysis for Italian Regions – Scenario *Temperature + Precipitation*

Chapter 7

Conclusions

This thesis has been a journey in the lands of complex systems and network science. We are now approaching the last mountain from where we can observe the walked path. From here, we can see in the distance the provinces of network analysis and inference, the realm of causality, the empire of probability and statistics, the kingdom of information theory, the green fields of environmental science, and at the foothills of the mountain the crowded roads of human mobility. Not least, on the shining horizon we descry the coastline of the sacred ocean of mystery. During this travel, we explored these and many other territories in which we learned as many ways of talking about humans on the move. In some cases we gave our contribution to the narration, which is gratifying, but most of all we are the ones who travelled all these regions transcending their boundaries to take a new perspective on the human mobility system. As a matter of fact, this was a journey to the pursuit of a fresh dip in that far away ocean. From there, we brought some treasures up here as keepsakes and placeholder for other wayfarer:

In the provinces of network analysis and inference we learned how to reconstruct a network of relations, which is useful to analyse the *shape* of the complex system that lead humans to move. In that ocean we found a new way to study this shape in a different manner, a "fuzzy" manner, that embrace the uncertainty about the relations and gives robust responses thanks to the tools acquired in the empire of probability. The world is hardly ever clear-cut, and so are the interconnections in complex systems such as the socio-economic and mobility systems. Usually, the fuzzy nature of these interconnections is refused and cut with thresholds, in order to recover a well defined, clear-cut, binary networked system. With our work we have provided a new method to study the topological characteristics of complex systems from a probabilistic perspective.

Subsequently, at the beginning of the COVID-19 pandemic, we crossed the fields of environmental science and we got closer to the region of human mobility. In that period we decided that as a young scientist I should gave my scientific

contribution in the debate about the pandemic and its consequences. Without loosing sight of complex systems and human mobility we drew a never-seen-before portrait of the network of relations between different flavors of human mobility data and other social and environmental variables. Moreover, we studied the effects of restrictions imposed on human mobility (and human activities in general) on the environmental system – notice that few months later we would have studied the opposite effects of environmental changes on human mobility... From time series measured by a satellite (Copernicus Sentinel-5P mission) we drawn out the moment in which the pandemic, and the consequent restrictions, began to have a measurable effect on the environment. To do so, we have (re)entered the kingdom of information-theory to define a novel information-theoretic measure of "dissimilarity" between time series observations, that served us to identify the tipping point in the system. From that point forward, we estimated a counterfactual reality, in which the restrictions had not been implemented, to reveal the difference with the factual reality of the lockdown. We found out that the behaviour of the environmental variable of interest were actually altered, but the fact that the restrictions were the cause of this alteration was not so sure. Additional data from the past and the future and further causal analysis may crack the mystery.

The journey continued with the study of data coming from dramatic events around the world. We investigated the consequences of catastrophic environmental events on human mobility. Natural/environmental disasters are indeed causes of human displacement and migration. We analysed data from different sources to understand the factors involved in shaping mobility patterns after tropical cyclones. We leveraged on the well-known gravity model to assess the association of human mobility with several factors, such as the social connectedness between origin and destination areas, the population living in those areas, the human development of those areas, the distance of those areas from each other and from the cyclones. Our analysis confirmed that the population density, the social and economic prosperity and the distance are useful predictors for mobility even in these particular cases, but our results are not complete nor definitive, since the data we held were not enough to have a sufficiently wide picture for a universal response.

The last leg of the journey was the Feature-Enriched Radiation Model (FERM), straight at the foothills of the mountain. We got on the shoulder of the giant *Radiation Model* and we guided him on a new way. We taught him how to handle information coming from the most diverse drivers of migrations. Specifically, we introduced the possibility to assign specific features to the geographic locations of origin and destination of mobility flows. These features serve to balance the influence of the population, that is the only feature allowed by the original model. The stochastic process is thus generalized to takes these features into account, but

the physical process of the early model remains the same. Therefore, the effect of intervening opportunities and population distribution can still influence the outcome. Indeed, assuming the features to be the same in all locations, the FERM can approximate the results of the Radiation Model with arbitrary precision. The calibration of the parameters is a non-trivial task, and a further effort is needed to find the correct optimization routine. Nevertheless, the simulation of the model under different scenarios gave optimistic results, that proved the virtues of the theoretical framework. To illustrate the behaviour of our FERM, we generated some "synthetic worlds" with a fractal distribution of locations that resembles the real world arrangement of cities and regions. We assigned a population to each location following the same distribution of real cities, and then we fixed their suitability features. The results were compared with the one of the original model and we discussed the new properties of our feature-enriched version. The FERM proved to be able to divert the mobility flows reflecting the arrangement of the features, without neglecting the global effects from the rest of the system. Accordingly, the flexibility of the model makes it suitable to adapt to real-world external conditions. For this reason, we wanted to test it on some real-world scenarios with real geographic locations and features. In particular, we applied the model to two case studies: Italian regions and USA countries. In both cases the features encoded in the model correspond to the climate suitability of different areas under different climate scenarios. The mobility networks obtained from this process suggest that the model works as expected, with some interesting additional properties. The model can reshape the mobility patterns at different scales, following the climatic drivers encoded in the features, and in addition it takes into account the global configuration of the system.

In conclusion, here we laid a methodological foundation for the study of complex systems from a probabilistic perspective and we proposed a new model to enhance the description and prediction accuracy of the human mobility system, once again assigning probabilities to mobility flows. The inherent complexity of the mobility system requires analytical tools robust against uncertainty at all levels. The probabilistic approach of our fuzzy model can handle the uncertainty and can be adopted to characterize the systems that contribute in causing the mobility patterns and may be used further to estimate the important features to be elected as descriptors of locations suitability. The probabilistic description of complex networks finds its seamless counterpart in the probability of the mobility fluxes modelled by the FERM, concluding the course of *structure, causes* and *effects*.

This stage of the journey is over, but the definitive and unifying frame for the causes, effects and patterns of human mobility is still blurred. Nevertheless, in this thesis we smoothed out some of the obstacles that pave the path for a

comprehensive mathematical model of human mobility. Such a model, requires to embrace the complexity of our world, from the uncertainty about causes and effects, to the chaotic evolution of the society and environment. We proposed some mathematical arguments that goes in this direction, and may open new insights to a thorough understanding of human mobility.

Bibliography

[1] Jared M Diamond and Doug Ordunio. Guns, germs, and steel. Books on Tape, 1999. 1

[2] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. nature, 453(7196):779–782, 2008. 1, 17

[3] Giorgio Fagiolo and Marina Mastrorillo. Does human migration affect international trade? a complex-network perspective. PloS one, 9(5):e97331, 2014. 1, 94

[4] Alessandro Galeazzi, Matteo Cinelli, Giovanni Bonaccorsi, Francesco Pierri, Ana Lucia Schmidt, Antonio Scala, Fabio Pammolli, and Walter Quattrociocchi. Human mobility in response to covid-19 in france, italy and uk. Scientific Reports, 11(1):1–10, 2021. 1

[5] Guy J Abel and Nikola Sander. Quantifying global international migration flows. Science, 343(6178):1520–1522, 2014. 1

[6] Nicholas Van Hear. Theories of migration and social change. Journal of ethnic and migration studies, 36(10):1531–1536, 2010. 1

[7] Layna Mosley and David A Singer. Migration, labor, and the international political economy. Annual Review of Political Science, 18:283–301, 2015. 1

[8] Moritz UG 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, et al. The effect of human mobility and control measures on the covid-19 epidemic in china. Science, 368(6490):493–497, 2020. 2, 94

[9] UN Doc. A/RES/73/195. Global compact for safe, orderly and regular migration (gcm). ., 2018. 2

[10] Richard Black, W Neil Adger, Nigel W Arnell, Stefan Dercon, Andrew Geddes, and David Thomas. The effect of environmental change on human migration. Global environmental change, 21:S3–S11, 2011. , 2, 6, 10, 39, 94

[11] Philip W Anderson. More is different. Science, 177(4047):393–396, 1972. 2

[12] Sebastian Raimondo and Manlio De Domenico. Measuring topological descriptors of complex networks under uncertainty. Physical Review E, 103(2):022311, 2021. 2, 41, 80

[13] Riccardo Gallotti, Francesco Valle, Nicola Castaldo, Pierluigi Sacco, and Manlio De Domenico. Assessing the risks of 'infodemics' in response to covid-19 epidemics. Nature Human Behaviour, 4(12):1285–1293, 2020. 3

[14] Nina Verstraete, Giuseppe Jurman, Giulia Bertagnolli, Arsham Ghavasieh, Vera Pancaldi, and Manlio De Domenico. Covmulnet19, integrating proteins, diseases, drugs, and symptoms: A network medicine approach to covid-19. Network and systems medicine, 3(1):130–141, 2020. 3

[15] Zahra Kolahchi, Manlio De Domenico, Lucina Q Uddin, Valentina Cauda, Igor Grossmann, Lucas Lacasa, Giulia Grancini, Morteza Mahmoudi, and Nima Rezaei. Covid-19 and its global economic impact. Advances in Experimental Medicine and Biology, 1318:825–837, 2021. 3

[16] Song Gao, Jinmeng Rao, Yuhao Kang, Yunlei Liang, and Jake Kruse. Mapping county-level mobility pattern changes in the united states in response to covid-19. SIGSpatial Special, 12(1):16–26, 2020. 3

[17] Mohammad Hassan Shakil, Ziaul Haque Munim, Mashiyat Tasnia, and Shahin Sarowar. Covid-19 and the environment: A critical review and research agenda. Science of the Total Environment, page 141022, 2020. 3

[18] Sebastian Raimondo, Barbara Benigni, and Manlio De Domenico. Covid-19 lockdown unravels the complex interplay between environmental conditions and human activity. Complexity, 2022. 3, 67

[19] 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. 3, 26, 27, 94, 96

[20] Gregoire Nicolis and Cathy Nicolis. Foundations of complex systems: emergence, information and predicition. World Scientific, 2012. 4

[21] Melanie Mitchell. Complexity: A guided tour. Oxford University Press, 2009. 4

[22] John H Holland. Hidden order: How adaptation builds complexity. Addison Wesley Longman Publishing Co., Inc., 1996. 4

[23] Murray Gell-Mann. The Quark and the Jaguar: Adventures in the Simple and the Complex. Macmillan, 1995. 4

[24] Cliff Hooker. Introduction to philosophy of complex systems: A: Part a: Towards a framework for complex systems. In Philosophy of complex systems, pages 3–90. Elsevier, 2011. 4

[25] James Ladyman and Karoline Wiesner. What is a complex system? Yale University Press, 2020. 4

[26] Joseph E. Brenner and Abir U. Igamberdiev. Structures and Complex Systems, pages 373–422. Springer International Publishing, Cham, 2021. 4

[27] Oriol Artime and Manlio De Domenico. From the origin of life to pandemics: emergent phenomena in complex systems, 2022. 4

[28] Peter A Corning. The re-emergence of "emergence": A venerable concept in search of a theory. Complexity, 7(6):18–30, 2002. 4, 5

[29] Sophie Gibb, Robin Findlay Hendry, and Tom Lancaster. The Routledge handbook of emergence. Routledge, 2019. 4, 5

[30] Kurt Koffka. Artikel „gestalt". Encyclopaedia of the Social Sciences, New York, 1931. 5

[31] Mark Bedau. Downward causation and the autonomy of weak emergence. Principia: an international journal of epistemology, 6(1):5–50, 2002. 5

[32] Erik P Hoel, Larissa Albantakis, and Giulio Tononi. Quantifying causal emergence shows that macro can beat micro. Proceedings of the National Academy of Sciences, 110(49):19790–19795, 2013. 5

[33] Flavio Del Santo and Nicolas Gisin. Physics without determinism: Alternative interpretations of classical physics. Physical Review A, 100(6):062107, 2019. 5

[34] Fernando E Rosas, Pedro AM Mediano, Henrik J Jensen, Anil K Seth, Adam B Barrett, Robin L Carhart-Harris, and Daniel Bor. Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data. PLOS Computational Biology, 16(12):e1008289, 2020. 5

[35] Sergio Chibbaro, Lamberto Rondoni, and Angelo Vulpiani. Reductionism, emergence and levels of reality. Springer, Berlin. by SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH (ETH) on, 3(20):17, 2014. 5

[36] Francis Heylighen, Paul Cilliers, and Carlos Gershenson. Complexity and philosophy. arXiv preprint cs/0604072, 2006. 5

[37] Orit Peleg, Jacob M Peters, Mary K Salcedo, and Lakshminarayanan Mahadevan. Collective mechanical adaptation of honeybee swarms. Nature Physics, 14(12):1193–1198, 2018. 5

[38] Colin W Clark and Marc Mangel. Foraging and flocking strategies: information in an uncertain environment. The American Naturalist, 123(5):626–641, 1984. 5

[39] Kolbjørn Tunstrøm, Yael Katz, Christos C Ioannou, Cristián Huepe, Matthew J Lutz, and Iain D Couzin. Collective states, multistability and transitional behavior in schooling fish. PLoS computational biology, 9(2):e1002915, 2013. 5

[40] Karl Friston. Life as we know it. Journal of the Royal Society Interface, 10(86):20130475, 2013. 5

[41] Steen Rasmussen, Liaohai Chen, David Deamer, David C Krakauer, Norman H Packard, Peter F Stadler, and Mark A Bedau. Transitions from nonliving to living matter. Science, 303(5660):963–965, 2004. 5

[42] Thomas C Schelling. Dynamic models of segregation. Journal of mathematical sociology, 1(2):143–186, 1971. 5

[43] Matteo Grasso, Larissa Albantakis, Jonathan P Lang, and Giulio Tononi. Causal reductionism and causal structures. Nature neuroscience, 24(10):1348–1355, 2021. 5, 28

[44] Stavros K Stavroglou, Athanasios A Pantelous, H Eugene Stanley, and Konstantin M Zuev. Unveiling causal interactions in complex systems. Proceedings of the National Academy of Sciences, 117(14):7599–7605, 2020. 5

[45] George FR Ellis. On the nature of causation in complex systems. Transactions of the Royal Society of South Africa, 63(1):69–84, 2008. 5

[46] Leyla Naghipour, Mohammad Taghi Aalami, and Vahid Nourani. Reconstruction of network connectivity by the interplay between complex structure and dynamics to discover climate networks. Theoretical and Applied Climatology, 143(3):969–987, 2021. 5

[47] George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. science, 338(6106):496–500, 2012. 5, 35, 37, 38, 45

[48] Stephen Castles and Mark J Miller. The migratory process and the formation of ethnic minorities. In The age of migration, pages 19–47. Springer, 1998. 6, 7, 8, 13

[49] International Organization for Migration. World Migration Report 2022. International Organization for Migration, 2021. 6, 39

[50] UN DESA. International migrant stock 2020. new york, 2021. 6

[51] Jaya Ramji-Nogales. Migration emergencies. Hastings LJ, 68:609, 2016. 6

[52] International Organization for Migration. World Migration Report 2020. International Organization for Migration, 2019. 6

[53] Ernest George Ravenstein. The laws of migration. Journal of the statistical society of London, 48(2):167–235, 1885. 6, 20, 94

[54] Ernest George Ravenstein. The laws of migration. Journal of the royal statistical society, 52(2):241–305, 1889. 6, 20, 94

[55] Douglas S Massey, Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J Edward Taylor. Theories of international migration: A review and appraisal. Population and development review, pages 431–466, 1993. 7

[56] Oded Stark and David E Bloom. The new economics of labor migration. The american Economic review, 75(2):173–178, 1985. 7

[57] Everett S Lee. A theory of migration. Demography, 3(1):47–57, 1966. 7

[58] Monica Boyd. Family and personal networks in international migration: recent developments and new agendas. International migration review, 23(3):638–670, 1989. 8

[59] Essam El-Hinnawi. Environmental refugees. UNEP, 1985. 8

[60] Astri Suhrke. Environmental degradation and population flows. Journal of International Affairs, pages 473–496, 1994. 8

[61] Hubert H Lamb. Climate, history and the modern world. Routledge, 2002. 8

[62] Katherine J Curtis, Elizabeth Fussell, and Jack DeWaard. Recovery migration after hurricanes katrina and rita: Spatial concentration and intensification in the migration system. Demography, 52(4):1269–1293, 2015. 8

[63] Xin Lu, Linus Bengtsson, and Petter Holme. Predictability of population displacement after the 2010 haiti earthquake. Proceedings of the National Academy of Sciences, 109(29):11576–11581, 2012. 8, 91

[64] Hanne Wiegel, Ingrid Boas, and Jeroen Warner. A mobilities perspective on migration in the context of environmental change. Wiley Interdisciplinary Reviews: Climate Change, 10(6):e610, 2019. 8

[65] Ingrid Boas, Carol Farbotko, Helen Adams, Harald Sterly, Simon Bush, Kees van der Geest, Hanne Wiegel, Hasan Ashraf, Andrew Baldwin, Giovanni Bettini, et al. Climate migration myths. Nature Climate Change, 9(12):901–903, 2019. 8, 13

[66] Clare Lizamit Samling. Environmental migration: A challenge for sustainable development. Environment, Development and Sustainability in India: Perspectives, Issues and Alternatives, page 165, 2021. 8

[67] Guy J Abel, Michael Brottrager, Jesus Crespo Cuaresma, and Raya Muttarak. Climate, conflict and forced migration. Global environmental change, 54:239–249, 2019. 8

[68] Jeremy Kendal, Jamshid J Tehrani, and John Odling-Smee. Human niche construction in interdisciplinary focus. Philosophical Transactions of the Royal Society B: Biological Sciences, 366(1566):785–792, 2011. 8

[69] Pratikshya Bohra-Mishra, Michael Oppenheimer, and Solomon M Hsiang. Nonlinear permanent migration response to climatic variations but minimal response to disasters. Proceedings of the National Academy of Sciences, 111(27):9780–9785, 2014. 9, 11

[70] Marshall Burke, Solomon M Hsiang, and Edward Miguel. Global non-linear effect of temperature on economic production. Nature, 527(7577):235–239, 2015. 9

[71] Robert McLeman. Thresholds in climate migration. Population and environment, 39(4):319–338, 2018. 9

[72] Chi Xu, Timothy A Kohler, Timothy M Lenton, Jens-Christian Svenning, and Marten Scheffer. Future of the human climate niche. Proceedings of the National Academy of Sciences, 117(21):11350–11355, 2020. , 9, 105, 106

[73] Jürgen Scheffran, Michael Brzoska, Jasmin Kominek, PMichael Link, Janpeter Schilling, et al. Climate change and violent conflict. Science(Washington), 336(6083):869–871, 2012. , 11

[74] Caleb Robinson, Bistra Dilkina, and Juan Moreno-Cruz. Modeling migration patterns in the usa under sea level rise. PloS One, 15(1):e0227436, 2020. 11

[75] Thomas F Homer-Dixon. Environmental scarcities and violent conflict: evidence from cases. International security, 19(1):5–40, 1994. 11

[76] Reto Knutti and Maria AA Rugenstein. Feedbacks, climate sensitivity and the limits of linear models. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373(2054):20150146, 2015. , 11, 12

[77] David D Zhang, Harry F Lee, Cong Wang, Baosheng Li, Qing Pei, Jane Zhang, and Yulun An. The causality analysis of climate change and large-scale human crisis. Proceedings of the National Academy of Sciences, 108(42):17296–17301, 2011. 11

[78] Lia Betti, Robert M Beyer, Eppie R Jones, Anders Eriksson, Francesca Tassi, Veronika Siska, Michela Leonardi, Pierpaolo Maisano Delser, Lily K Bentley, Philip R Nigst, et al. Climate shaped how neolithic farmers and european hunter-gatherers interacted after a major slowdown from 6,100 bce to 4,500 bce. Nature Human Behaviour, 4(10):1004–1010, 2020. 11

[79] Sabine L Perch-Nielsen, Michèle B Bättig, and Dieter Imboden. Exploring the link between climate change and migration. Climatic change, 91(3-4):375–393, 2008. 13

[80] Richard Black, Nigel W Arnell, W Neil Adger, David Thomas, and Andrew Geddes. Migration, immobility and displacement outcomes following extreme events. Environmental Science & Policy, 27:S32–S43, 2013. 13, 88

[81] Diane C Bates. Environmental refugees? classifying human migrations caused by environmental change. Population and environment, 23(5):465–477, 2002. 13

[82] Christine Gibb and James Ford. Should the united nations framework convention on climate change recognize climate migrants? Environmental Research Letters, 7(4):045601, 2012. 13

[83] Karen Elizabeth McNamara. Conceptualizing discourses on environmental refugees at the united nations. Population and Environment, 29(1):12–24, 2007. 13

[84] Simon Behrman and Avidan Kent. Climate refugees: beyond the legal impasse? Routledge, 2018. 13

[85] Richard Black, Stephen RG Bennett, Sandy M Thomas, and John R Beddington. Migration as adaptation. Nature, 478(7370):447–449, 2011. 13

[86] Melissa Haeffner, Jacopo A Baggio, and Kathleen Galvin. Investigating environmental migration and other rural drought adaptation strategies in baja california sur, mexico. Regional Environmental Change, 18(5):1495–1507, 2018. 13

[87] Elena Ambrosetti and Enza Roberta Petrillo. Environmental disasters, migration and displacement. insights and developments from l'aquila's case. Environmental Science & Policy, 56:80–88, 2016. 13

[88] Clark L Gray and Valerie Mueller. Natural disasters and population mobility in bangladesh. Proceedings of the National Academy of Sciences, 109(16):6000–6005, 2012. 13

[89] Karl Pearson. The problem of the random walk. Nature, 72(1865):294–294, 1905. 15

[90] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. Stochastic processes, volume 2. Wiley New York, 1996. 16

[91] Dirk Brockmann, Lars Hufnagel, and Theo Geisel. The scaling laws of human travel. Nature, 439(7075):462–465, 2006. 17

[92] Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. Modelling the scaling properties of human mobility. Nature physics, 6(10):818–823, 2010. 17

[93] Michael S Salkin, Theodore P Lianos, and Quirino Paris. Population predictions for the western united states: a markov chain approach. Journal of Regional Science, 15(1):53–60, 1975. 17

[94] Jie Pan and Anna Nagurney. Using markov chains to model human migration in a network equilibrium framework. Mathematical and computer modelling, 19(11):31–39, 1994. 17

[95] Hyunuk Kim and Ha Yoon Song. Formulating human mobility model in a form of continuous time markov chain. Procedia Computer Science, 10:389–396, 2012. 17

[96] Vincent Huang and James Unwin. Markov chain models of refugee migration data. IMA Journal of Applied Mathematics, 85(6):892–912, 2020. 17

[97] James R Norris and James Robert Norris. Markov chains. Cambridge university press, 1998. 17

[98] Volker Grimm, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara Grand, Simone K Heinz, Geir Huse, et al. A standard protocol for describing individual-based and agent-based models. Ecological modelling, 198(1-2):115–126, 2006. 18

[99] Uri Wilensky and William Rand. An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo. Mit Press, 2015. 18

[100] Sameera Abar, Georgios K Theodoropoulos, Pierre Lemarinier, and Gregory MP O'Hare. Agent based modelling and simulation tools: A review of the state-of-art software. Computer Science Review, 24:13–33, 2017. 18

[101] Anna Klabunde and Frans Willekens. Decision-making in agent-based models of migration: state of the art and challenges. European Journal of Population, 32(1):73–97, 2016. 18

[102] Robert L Axtell. Coordination in transient social networks: An agent-based computational model of the timing of retirement robert l. axtell and joshua m. epstein. Generative social science: Studies in agent-based computational modeling, 146, 2006. 18

[103] Frank Schweitzer. Modelling migration and economic agglomeration with active brownian particles. Adv. Complex Syst., 1:11–38, 1998. 18

[104] Ruohong Cai and Michael Oppenheimer. An agent-based model of climate-induced agricultural labor migration. Research in Agricultural & Applied Economics, 2013. 18

[105] Francesco C Billari, Thomas Fent, Alexia Prskawetz, and Jürgen Scheffran. Agent-based computational modelling: an introduction. In Agent-based computational modelling, pages 1–16. Springer, 2006. 18

[106] Dominic Kniveton, Christopher Smith, and Sharon Wood. Agent-based model simulations of future changes in migration flows for burkina faso. Global Environmental Change, 21:S34–S40, 2011. 18

[107] Anna Klabunde. Computational economic modeling of migration. Available at SSRN 2470525, 2014. 18

[108] Diana Suleimenova, David Bell, and Derek Groen. A generalized simulation development approach for predicting refugee destinations. Scientific reports, 7(1):1–13, 2017. 18

[109] Corinna Elsenbroich. Explanation in agent-based modelling: Functions, causality or mechanisms? JASSS, 15(3), 2012. 18

[110] Kenneth E Train. Discrete choice methods with simulation. Cambridge university press, 2009. 19, 20

[111] Daniel McFadden. The measurement of urban travel demand. Journal of public economics, 3(4):303–328, 1974. 20

[112] Simone Bertoli and Jesús Fernández-Huertas Moraga. Multilateral resistance to migration. Journal of development economics, 102:79–100, 2013. 20, 24

[113] George Kingsley Zipf. The p 1 p 2/d hypothesis: on the intercity movement of persons. American sociological review, 11(6):677–686, 1946. 20

[114] Henry Charles Carey. Principles of social science, volume 3. JB Lippincott & Company, 1871. 20

[115] William John Reilly. The law of retail gravitation. WJ Reilly, 1931. 20

[116] Ernest Charles Young. The movement of farm population, volume 426. Cornell University Agricultural Experiment Station, 1924. 20

[117] Jan Tinbergen. Shaping the world economy; suggestions for an international economic policy. Periodicals Service Co, 1962. 21, 94

[118] Alan Wilson. A statistical theory of spatial distribution models. Transportation Research, 1:253–269, 1967. 22

[119] Oleguer Sagarra, CJ Pérez Vicente, and Albert Dïaz-Guilera. Statistical mechanics of multiedge networks. Physical Review E, 88(6):062806, 2013. 22, 23

[120] O Sagarra, CJ Pérez Vicente, and A Díaz-Guilera. Role of adjacency-matrix degeneracy in maximum-entropy-weighted network models. Physical Review E, 92(5):052816, 2015. 23

[121] Jacques Poot, Omoniyi Alimi, Michael P Cameron, and David C Maré. The gravity model of migration: the successful comeback of an ageing superstar in regional science. Available at SSRN 2864830, 2016. 23, 89

[122] Adolfo Maza, María Gutiérrez-Portilla, María Hierro, and José Villaverde. Internal migration in spain: Dealing with multilateral resistance and nonlinearites. International Migration, 57(1):75–93, 2019. 23

[123] Gautier Krings, Francesco Calabrese, Carlo Ratti, and Vincent D Blondel. Urban gravity: a model for inter-city telecommunication flows. Journal of Statistical Mechanics: Theory and Experiment, 2009(07):L07003, 2009. 23

[124] 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. 23

[125] Cécile Viboud, Ottar N Bjørnstad, David L Smith, Lone Simonsen, Mark A Miller, and Bryan T Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. science, 312(5772):447–451, 2006. 23

[126] JMC Santos Silva and Silvana Tenreyro. The log of gravity. The Review of Economics and statistics, 88(4):641–658, 2006. 23

[127] Peter McCullagh and John A Nelder. Generalized linear models. Routledge, 2019. 23

[128] A Colin Cameron and Pravin K Trivedi. Essentials of count data regression. A companion to theoretical econometrics, 331, 2001. 23

[129] Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. In Handbook of international economics, volume 4, pages 131–195. Elsevier, 2014. 23, 24

[130] Anna-Lena Wölwer, Martin Breßlein, and Jan Pablo Burgard. Gravity models in r. Austrian Journal of Statistics, 47(4):16–35, 2018. 23

[131] Ben Shepherd, Hrisyana Doytchinova, Alexey Kravchenko, et al. The gravity model of international trade: a user guide [r vision], 2019. 23

[132] James E Anderson. A theoretical foundation for the gravity equation. The American economic review, 69(1):106–116, 1979. 23

[133] James E Anderson and Eric Van Wincoop. Gravity with gravitas: A solution to the border puzzle. American economic review, 93(1):170–192, 2003. 23

[134] Robert C Feenstra. Advanced international trade: theory and evidence. Princeton university press, 2015. 24

[135] Stephen Redding and Anthony J Venables. Economic geography and international inequality. Journal of international Economics, 62(1):53–82, 2004. 24

[136] Filippo Simini, Gianni Barlacchi, Massimilano Luca, and Luca Pappalardo. A deep gravity model for mobility flows generation. Nature communications, 12(1):1–13, 2021. 24

[137] Samuel A Stouffer. Intervening opportunities: a theory relating mobility and distance. American sociological review, 5(6):845–867, 1940. 24, 94

[138] Stanislas Dehaene. The neural basis of the weber–fechner law: a logarithmic mental number line. Trends in cognitive sciences, 7(4):145–147, 2003. 25

[139] Morton Schneider. Gravity models and trip distribution theory. Papers in Regional Science, 5(1):51–56, 1959. 25, 94

[140] John T Lynch, Glenn E Brokke, Alan M Voorhees, and Morton Schneider. Panel discussion on inter-area travel formulas. National Academy of Science-National Council, Highway Research Board, Bulletin, 253:128–138, 1960. 25

[141] David K Witheford. Comparison of trip distribution by opportunity model and gravity model. Pittsburgh Area Transportation Study, 1961. 26

[142] Howard C Lawson and John A Dearinger. A comparison of four work trip distribution models. Journal of the Highway Division, 93(2):1–25, 1967. 26

[143] F Zhao, LF Chow, MT Li, A Gan, and DL Shen. Refinement of fsutms trip distribution methodology. technical memorandum no. 3: Calibration of an intervening opportunity model for palm beach county. Technical report, National Academy of Sciences, Enginnering and Medicine, 2001. 26

[144] Alan Geoffrey Wilson. Urban and regional models in geography and planning. John Wiley & Sons Ltd, 1974. 26

[145] Kingsley E Haynes, Dudley L Poston Jr, and Paul Schnirring. Intermetropolitan migration in high and low opportunity areas: indirect tests of the distance and intervening opportunities hypotheses. Economic Geography, 49(1):68–73, 1973. 26

[146] Earl R Ruiter. Toward a better understanding of the intervening opportunities model. Transportation Research, 1(1):47–56, 1967. 26

[147] Timothy J Fik and Gordon F Mulligan. Spatial flows and competing central places: towards a general theory of hierarchical interaction. Environment and Planning A, 22(4):527–549, 1990. 26

[148] Amy Wesolowski, Wendy Prudhomme O'Meara, Nathan Eagle, Andrew J Tatem, and Caroline O Buckee. Evaluating spatial interaction models for regional mobility in sub-saharan africa. PLoS computational biology, 11(7):e1004267, 2015. 27

[149] Emmanouil Tranos, Masood Gheasi, and Peter Nijkamp. International migration: a global complex network. Environment and Planning B: Planning and Design, 42(1):4–22, 2015. 28

[150] Kyle F Davis, Paolo D'Odorico, Francesco Laio, and Luca Ridolfi. Global spatio-temporal patterns in human migration: a complex network perspective. PloS one, 8(1):e53723, 2013. 28

[151] Michael Windzio. The network of global migration 1990–2013: Using ergms to test theories of migration between countries. Social Networks, 53:20–29, 2018. 28

[152] Paolo Sgrignoli, Rodolfo Metulini, Stefano Schiavo, and Massimo Riccaboni. The relation between global migration and trade networks. Physica A: Statistical Mechanics and its Applications, 417:245–260, 2015. 28

[153] Alexander Belyi, Iva Bojic, Stanislav Sobolevsky, Izabela Sitko, Bartosz Hawelka, Lada Rudikova, Alexander Kurbatski, and Carlo Ratti. Global multi-layer network of human mobility. International Journal of Geographical Information Science, 31(7):1381–1402, 2017. 28

[154] Mario Bunge. Causality and modern science. Routledge, 2017. 28

[155] Phyllis Illari and Federica Russo. Causality: Philosophical theory meets scientific practice. OUP Oxford, 2014. 28

[156] Judea Pearl and Dana Mackenzie. The book of why: the new science of cause and effect. Basic books, 2018. 28, 29

[157] Phillip H Delacy. The problem of causation in plato's philosophy. Classical Philology, 34(2):97–115, 1939. 28

[158] Karl Pearson. The Grammar of Science. Adam and Charles Black, London, 1911. 28

[159] David Hume. A treatise of human nature. Clarendon Press, 1896. 28

[160] Carl Gustav Jung. La sincronicità. Bollati Boringhieri, 2021. 28

[161] Judea Pearl. Causality. Cambridge university press, 2009. 28, 29, 31

[162] Marco Baldovin, Fabio Cecconi, and Angelo Vulpiani. Understanding causation via correlations and linear response theory. Physical Review Research, 2(4):043436, 2020. 28

[163] Klaas Enno Stephan, Will D Penny, Rosalyn J Moran, Hanneke EM den Ouden, Jean Daunizeau, and Karl J Friston. Ten simple rules for dynamic causal modeling. Neuroimage, 49(4):3099–3109, 2010. 29

[164] Steven L Scott and Hal R Varian. Predicting the present with bayesian structural time series. International Journal of Mathematical Modelling and Numerical Optimisation, 5(1-2):4–23, 2014. 29, 82

[165] Kay H Brodersen, Fabian Gallusser, Jim Koehler, Nicolas Remy, Steven L Scott, et al. Inferring causal impact using bayesian structural time-series models. The Annals of Applied Statistics, 9(1):247–274, 2015. 29, 71, 82, 83

[166] Andrew Gelman and Jennifer Hill. Data analysis using regression and multilevel/hierarchical models. Cambridge university press, 2006. 29

[167] Avi Feller and Andrew Gelman. Hierarchical models for causal effects. Emerging Trends in the Social and Behavioral Sciences: An interdisciplinary, searchable, and linkable resource, pages 1–16, 2015. 29

[168] James M McCracken. Exploratory causal analysis with time series data. Synthesis Lectures on Data Mining and Knowledge Discovery, 8(1):1–147, 2016. 29

[169] Mikail Rubinov and Olaf Sporns. Weight-conserving characterization of complex functional brain networks. Neuroimage, 56(4):2068–2079, 2011. 29

[170] Takayuki Mizuno, Hideki Takayasu, and Misako Takayasu. Correlation networks among currencies. Physica A: Statistical Mechanics and its Applications, 364:336–342, 2006. 29

[171] Michael M Saint-Antoine and Abhyudai Singh. Network inference in systems biology: recent developments, challenges, and applications. Current opinion in biotechnology, 63:89–98, 2020. 29

[172] Analyse Mathematique. Sur les probabilités des erreurs de situation d'un point mem. Acad. Roy. Sci. Inst. France, Sci. Math, et Phys, 9:255–332, 1844. 29

[173] Francis Galton. I. co-relations and their measurement, chiefly from anthropometric data. Proceedings of the Royal Society of London, 45(273-279):135–145, 1889. 29

[174] Karl Pearson. Vii. note on regression and inheritance in the case of two parents. proceedings of the royal society of London, 58(347-352):240–242, 1895. 29

[175] Clive WJ Granger. Testing for causality: A personal viewpoint. Journal of Economic Dynamics and control, 2:329–352, 1980. 30, 32

[176] Danielle S Bassett, Mason A Porter, Nicholas F Wymbs, Scott T Grafton, Jean M Carlson, and Peter J Mucha. Robust detection of dynamic community structure in networks. Chaos: An Interdisciplinary Journal of Nonlinear Science, 23(1):013142, 2013. 31

[177] Hans Von Storch and Francis W Zwiers. Statistical analysis in climate research. Cambridge university press, 2002. 31

[178] Andrew Zalesky, Alex Fornito, and Ed Bullmore. On the use of correlation as a measure of network connectivity. Neuroimage, 60(4):2096–2106, 2012. 31

[179] Katerina Schindler-Hlavackova, Milan Palus, Martin Vejmelka, and Joydeep Bhattacharya. Causality detection based on information-theoretic approach in time series analysis. Phys. Rep, 441:1–46, 2007. 31

[180] Tristan Millington and Mahesan Niranjan. Partial correlation financial networks. Applied Network Science, 5(1):1–19, 2020. 31

[181] Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society, pages 424–438, 1969. 32, 71, 81

[182] Bethany Lusch, Pedro D Maia, and J Nathan Kutz. Inferring connectivity in networked dynamical systems: Challenges using granger causality. Physical Review E, 94(3):032220, 2016. 33

[183] Luca Faes, Giandomenico Nollo, Sebastiano Stramaglia, and Daniele Marinazzo. Multiscale granger causality. Physical Review E, 96(4):042150, 2017. 33

[184] Shuixia Guo, Anil K Seth, Keith M Kendrick, Cong Zhou, and Jianfeng Feng. Partial granger causality—eliminating exogenous inputs and latent variables. Journal of neuroscience methods, 172(1):79–93, 2008. 33

[185] Ralf Steuer, Jürgen Kurths, Carsten O Daub, Janko Weise, and Joachim Selbig. The mutual information: detecting and evaluating dependencies between variables. Bioinformatics, 18(suppl_2):S231–S240, 2002. 34

[186] Stefan Frenzel and Bernd Pompe. Partial mutual information for coupling analysis of multivariate time series. Physical review letters, 99(20):204101, 2007. 34

[187] Juan Zhao, Yiwei Zhou, Xiujun Zhang, and Luonan Chen. Part mutual information for quantifying direct associations in networks. Proceedings of the National Academy of Sciences, 113(18):5130–5135, 2016. 34

[188] Thomas Schreiber. Measuring information transfer. Physical review letters, 85(2):461, 2000. 34, 35

[189] Lionel Barnett, Adam B Barrett, and Anil K Seth. Granger causality and transfer entropy are equivalent for gaussian variables. Physical review letters, 103(23):238701, 2009. 35

[190] Floris Takens. Detecting strange attractors in turbulence. In Dynamical systems and turbulence, Warwick 1980, pages 366–381. Springer, 1981. 35, 36

[191] Terry Bossomaier, Lionel Barnett, Michael Harré, and Joseph T Lizier. Transfer entropy. In An introduction to transfer entropy, pages 65–95. Springer, 2016. 35

[192] Joseph T Lizier and Mikhail Prokopenko. Differentiating information transfer and causal effect. The European Physical Journal B, 73(4):605–615, 2010. 35

[193] Hassler Whitney. Differentiable manifolds. Annals of Mathematics, pages 645–680, 1936. 37

[194] Tim Sauer, James A Yorke, and Martin Casdagli. Embedology. Journal of statistical Physics, 65(3):579–616, 1991. 37

[195] Valeria d'Andrea and Manlio De Domenico. Compressing phase space detects state changes in nonlinear dynamical systems. Complexity, 2020, 2020. 37

[196] Hong-guang Ma and Chong-zhao Han. Selection of embedding dimension and delay time in phase space reconstruction. Frontiers of Electrical and Electronic Engineering in China, 1(1):111–114, 2006. 37

[197] Michael Small and Chi Kong Tse. Optimal embedding parameters: a modelling paradigm. Physica D: Nonlinear Phenomena, 194(3-4):283–296, 2004. 37

[198] M Ataei, A Khaki-Sedigh, B Lohmann, and C Lucas. Determination of embedding dimension using multiple time series based on singular value decomposition. In Proceedings of the Fourth International Symposium on Mathematical Modeling. Vienna, Austria, pages 190–196, 2003. 37

[199] Matthew B Kennel, Reggie Brown, and Henry DI Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physical review A, 45(6):3403, 1992. 37

[200] Huanfei Ma, Kazuyuki Aihara, and Luonan Chen. Detecting causality from nonlinear dynamics with short-term time series. Scientific reports, 4(1):1–10, 2014. 38

[201] Adam Thomas Clark, Hao Ye, Forest Isbell, Ethan R Deyle, Jane Cowles, G David Tilman, and George Sugihara. Spatial convergent cross mapping to detect causal relationships from short time series. Ecology, 96(5):1174–1181, 2015. 38

[202] James M McCracken and Robert S Weigel. Convergent cross-mapping and pairwise asymmetric inference. Physical Review E, 90(6):062903, 2014. 38, 43

[203] Carlos Gershenson. The implications of interactions for science and philosophy. Foundations of Science, 18(4):781–790, 2013. 39

[204] Mark EJ Newman. The structure and function of complex networks. SIAM review, 45(2):167–256, 2003. 42

[205] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. Physics reports, 424(4-5):175–308, 2006. 42

[206] Yue Yang and Huijie Yang. Complex network-based time series analysis. Physica A: Statistical Mechanics and its Applications, 387(5):1381 – 1386, 2008. 42

[207] Yong Zou, Reik V Donner, Norbert Marwan, Jonathan F Donges, and Jürgen Kurths. Complex network approaches to nonlinear time series analysis. Physics Reports, 787:1–97, 2019. 42

[208] Manlio De Domenico, Shuntaro Sasai, and Alex Arenas. Mapping multiplex hubs in human functional brain networks. Frontiers in neuroscience, 10:326, 2016. 42

[209] M. Chavez, M. Valencia, V. Navarro, V. Latora, and J. Martinerie. Functional modularity of background activities in normal and epileptic brain networks. Phys. Rev. Lett., 104:118701, Mar 2010. 42

[210] Volker Pernice, Benjamin Staude, Stefano Cardanobile, and Stefan Rotter. How structure determines correlations in neuronal networks. PLoS computational biology, 7(5), 2011. 42

[211] Danielle S Bassett and Olaf Sporns. Network neuroscience. Nature neuroscience, 20(3):353, 2017. 42

[212] Sunil Kumar and Nivedita Deo. Correlation and network analysis of global financial indices. Physical Review E, 86(2):026101, 2012. 42, 43, 68

[213] Giovanni Bonanno, Guido Caldarelli, Fabrizio Lillo, and Rosario N Mantegna. Topology of correlation-based minimal spanning trees in real and model markets. Physical Review E, 68(4):046130, 2003. 42, 45

[214] Anastasios A Tsonis and Kyle L Swanson. Topology and predictability of el nino and la nina networks. Physical Review Letters, 100(22):228502, 2008. 42

[215] Dong Zhou, Avi Gozolchiani, Yosef Ashkenazy, and Shlomo Havlin. Teleconnection paths via climate network direct link detection. Physical review letters, 115(26):268501, 2015. 42

[216] Bellie Sivakumar and Fitsum M Woldemeskel. A network-based analysis of spatial rainfall connections. Environmental Modelling & Software, 69:55–62, 2015. 42

[217] Niklas Boers, Bedartha Goswami, Aljoscha Rheinwalt, Bodo Bookhagen, Brian Hoskins, and Jürgen Kurths. Complex networks reveal global pattern of extreme-rainfall teleconnections. Nature, 566(7744):373–377, 2019. 42, 45

[218] Réka Albert. Network inference, analysis, and modeling in systems biology. The Plant Cell, 19(11):3327–3338, 2007. 42

[219] Michael Hecker, Sandro Lambeck, Susanne Toepfer, Eugene Van Someren, and Reinhard Guthke. Gene regulatory network inference: data integration in dynamic models—a review. Biosystems, 96(1):86–103, 2009. 42

[220] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience, 10(3):186–198, 2009. 42

[221] Brian J Hoskins and David J Karoly. The steady linear response of a spherical atmosphere to thermal and orographic forcing. Journal of the Atmospheric Sciences, 38(6):1179–1196, 1981. 43

[222] Jose Casadiego and Marc Timme. Network dynamics as an inverse problem. In Mathematical technology of networks, pages 39–48. Springer, 2015. 43

[223] Mor Nitzan, Jose Casadiego, and Marc Timme. Revealing physical interaction networks from statistics of collective dynamics. Science advances, 3(2):e1600396, 2017. 43

[224] Lucas Lacasa, Bartolo Luque, Fernando Ballesteros, Jordi Luque, and Juan Carlos Nuño. From time series to complex networks: The visibility graph. Proceedings of the National Academy of Sciences, 105(13):4972–4975, 2008. 43

[225] Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11):eaau4996, 2019. 43, 44, 74

[226] Jakob Runge, Jobst Heitzig, Norbert Marwan, and Jürgen Kurths. Quantifying causal coupling strength: A lag-specific measure for multivariate time series related to transfer entropy. Phys. Rev. E, 86:061121, Dec 2012. 43

[227] Jakob Runge, Sebastian Bathiany, Erik Bollt, Gustau Camps-Valls, Dim Coumou, Ethan Deyle, Clark Glymour, Marlene Kretschmer, Miguel D Mahecha, Jordi Muñoz-Marí, et al. Inferring causation from time series in earth system sciences. Nature communications, 10(1):1–13, 2019. 43

[228] Marlene Kretschmer, Dim Coumou, Jonathan F Donges, and Jakob Runge. Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation. Journal of Climate, 29(11):4069–4081, 2016. 43

[229] Marc Timme and Jose Casadiego. Revealing networks from dynamics: an introduction. Journal of Physics A: Mathematical and Theoretical, 47(34):343001, 2014. 43, 68

[230] Massimiliano Zanin, Seddik Belkoura, Javier Gomez, César Alfaro, and Javier Cano. Topological structures are consistently overestimated in functional complex networks. Scientific reports, 8(1):1–9, 2018. 43

[231] Marc Timme. Revealing network connectivity from response dynamics. Phys. Rev. Lett., 98:224101, May 2007. 43

[232] Mary-Ellen Lynall, Danielle S Bassett, Robert Kerwin, Peter J McKenna, Manfred Kitzbichler, Ulrich Muller, and Ed Bullmore. Functional connectivity and brain networks in schizophrenia. Journal of Neuroscience, 30(28):9477–9487, 2010. 43

[233] Jie Sun, Dane Taylor, and Erik M. Bollt. Causal network inference by optimal causation entropy. SIAM Journal on Applied Dynamical Systems, 14(1):73–106, 2015. 43

[234] Jakob Runge, Jobst Heitzig, Vladimir Petoukhov, and Jürgen Kurths. Escaping the curse of dimensionality in estimating multivariate transfer entropy. Phys. Rev. Lett., 108:258701, Jun 2012. 43

[235] Danielle S Bassett, Nicholas F Wymbs, Mason A Porter, Peter J Mucha, and Scott T Grafton. Cross-linked structure of network evolution. Chaos: An Interdisciplinary Journal of Nonlinear Science, 24(1):013112, 2014. 43, 80

[236] Michael M Saint-Antoine and Abhyudai Singh. Network inference in systems biology: recent developments, challenges, and applications. Current Opinion in Biotechnology, 63:89–98, 2020. 43

[237] František Váša, Edward T Bullmore, and Ameera X Patel. Probabilistic thresholding of functional connectomes: Application to schizophrenia. Neuroimage, 172:326–340, 2018. 43

[238] George T. Cantwell, Yanchen Liu, Benjamin F. Maier, Alice C. Schwarze, Carlos A. Serván, Jordan Snyder, and Guillaume St-Onge. Thresholding normally distributed data creates complex networks. Phys. Rev. E, 101:062302, Jun 2020. 43

[239] Tomomichi Nakamura, Toshihiro Tanizawa, and Michael Small. Constructing networks from a dynamical system perspective for multivariate nonlinear time series. Physical Review E, 93(3):032323, 2016. 44, 46, 68, 74, 80

[240] Sandrine Dudoit, Juliet Popper Shaffer, and Jennifer C Boldrick. Multiple hypothesis testing in microarray experiments. Statistical Science, pages 71–103, 2003. 44

[241] Tiago P Peixoto. Network reconstruction and community detection from dynamics. Physical review letters, 123(12):128301, 2019. 44, 45, 68, 74

[242] Jean-Gabriel Young, George T. Cantwell, and M. E. J. Newman. Robust bayesian inference of network structure from unreliable data, 2016. 44

[243] Peer Bork, Lars J Jensen, Christian von Mering, Arun K Ramani, Insuk Lee, and Edward M Marcotte. Protein interaction networks from yeast to human. Current Opinion in Structural Biology, 14(3):292 – 299, 2004. 44

[244] Sicheng Dai, Hélène Bouchet, Aurélie Nardy, Eric Fleury, Jean-Pierre Chevrot, and Márton Karsai. Temporal social network reconstruction using wireless proximity sensors: model selection and consequences. EPJ Data Science, 9(1):19, 2020. 44

[245] Daniel Chen, Leonidas J Guibas, John Hershberger, and Jian Sun. Road network reconstruction for organizing paths. In Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms, pages 1309–1320. SIAM, 2010. 44

[246] Till Hoffmann, Leto Peel, Renaud Lambiotte, and Nick S Jones. Community detection in networks without observing edges. Science Advances, 6(4):eaav1478, 2020. 44

[247] Tiago P Peixoto. Bayesian stochastic blockmodeling. Advances in network clustering and blockmodeling, pages 289–332, 2019. 45

[248] L Mandel and E Wolf. Spectral coherence and the concept of cross-spectral purity. JOSA, 66(6):529–535, 1976. 45

[249] Claude E Shannon and Warren Weaver. The mathematical theory of information. Urbana: University of Illinois Press, 97, 1949. 45

[250] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience, 10(3):186–198, 2009. 45

[251] Axel Wismüller, Xixi Wang, Adora M DSouza, and Mahesh B Nagarajan. A framework for exploring non-linear functional connectivity and causality in

the human brain: mutual connectivity analysis (mca) of resting-state functional mri with convergent cross-mapping and non-metric clustering. arXiv preprint arXiv:1407.3809, 2014. 45

[252] Jaeseung Jeong, John C Gore, and Bradley S Peterson. Mutual information analysis of the eeg in patients with alzheimer's disease. Clinical neurophysiology, 112(5):827–835, 2001. 45

[253] Jonathan Schiefer, Alexander Niederbühl, Volker Pernice, Carolin Lennartz, Jürgen Hennig, Pierre LeVan, and Stefan Rotter. From correlation to causation: Estimating effective connectivity from zero-lag covariances of brain signals. PLoS computational biology, 14(3):e1006056, 2018. 45

[254] Jonathan F Donges, Yong Zou, Norbert Marwan, and Jürgen Kurths. Complex networks in climate dynamics. The European Physical Journal Special Topics, 174(1):157–179, 2009. 45

[255] Jonathan F Donges, Yong Zou, Norbert Marwan, and Jürgen Kurths. The backbone of the climate network. EPL (Europhysics Letters), 87(4):48007, 2009. 45, 68

[256] Kazuko Yamasaki, Avi Gozolchiani, and Shlomo Havlin. Climate networks around the globe are significantly affected by el niño. Physical review letters, 100(22):228501, 2008. 45

[257] A Namaki, AH Shirazi, R Raei, and GR Jafari. Network analysis of a financial market based on genuine correlation and threshold method. Physica A: Statistical Mechanics and its Applications, 390(21-22):3835–3841, 2011. 45

[258] Thomas Sellke, MJ Bayarri, and James O Berger. Calibration of ρ values for testing precise null hypotheses. The American Statistician, 55(1):62–71, 2001. 46, 47, 62

[259] Leonhard Held. A nomogram for pvalues. BMC medical research methodology, 10(1):21, 2010. 46

[260] HM James Hung, Robert T O'Neill, Peter Bauer, and Karl Kohne. The behavior of the p-value when the alternative hypothesis is true. Biometrics, pages 11–22, 1997. 46

[261] Patrick Billingsley. Probability and measure. John Wiley & Sons, 2008. 50

[262] Mark Newman. Networks. Oxford university press, 2018. 52

[263] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. nature, 393(6684):440, 1998. 52

[264] Pablo M Gleiser and Leon Danon. Community structure in jazz. Advances in complex systems, 6(04):565–573, 2003. 54

[265] Neo D Martinez. Artifacts or attributes? effects of resolution on the little rock lake food web. Ecological monographs, 61(4):367–392, 1991. 54

[266] Logan Harriger, Martijn P Van Den Heuvel, and Olaf Sporns. Rich club organization of macaque cerebral cortex and its role in network communication. PloS one, 7(9), 2012. 54

[267] Robert M May. Simple mathematical models with very complicated dynamics. Nature, 261(5560):459–467, 1976. 54

[268] Gemma Lancaster, Dmytro Iatsenko, Aleksandra Pidde, Valentina Ticcinelli, and Aneta Stefanovska. Surrogate data for hypothesis testing of physical systems. Physics Reports, 748:1–60, 2018. 54

[269] Yili Hong. On computing the distribution function for the poisson binomial distribution. Computational Statistics & Data Analysis, 59:41–51, 2013. 56

[270] Bhavesh R Borate, Elissa J Chesler, Michael A Langston, Arnold M Saxton, and Brynn H Voy. Comparison of threshold selection methods for microarray gene co-expression matrices. BMC research notes, 2(1):240, 2009. 56

[271] Fabrizio De Vico Fallani, Vito Latora, and Mario Chavez. A topological criterion for filtering information in complex brain networks. PLoS computational biology, 13(1):e1005305, 2017. 56

[272] Vittoria Colizza, Alessandro Flammini, M Angeles Serrano, and Alessandro Vespignani. Detecting rich-club ordering in complex networks. Nature physics, 2(2):110–115, 2006. 59

[273] Duncan J Murdoch, Yu-Ling Tsai, and James Adcock. P-values are random variables. The American Statistician, 62(3):242–245, 2008. 62

[274] Manlio De Domenico, Albert Solé-Ribalta, Emanuele Cozzo, Mikko Kivelä, Yamir Moreno, Mason A Porter, Sergio Gómez, and Alex Arenas. Mathematical formulation of multilayer networks. Physical Review X, 3(4):041022, 2013. 68

[275] Jianxi Gao, Sergey V Buldyrev, H Eugene Stanley, and Shlomo Havlin. Networks formed from interdependent networks. Nature physics, 8(1):40–48, 2012. 68

[276] Sergey V Buldyrev, Roni Parshani, Gerald Paul, H Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. Nature, 464(7291):1025–1028, 2010. 68

[277] Amir Bashan, Yehiel Berezin, Sergey V Buldyrev, and Shlomo Havlin. The extreme vulnerability of interdependent spatially embedded networks. Nature Physics, 9(10):667–672, 2013. 68

[278] Alessandro Vespignani. The fragility of interdependency. Nature, 464(7291):984–985, 2010. 68

[279] Jose Casadiego and Marc Timme. Network dynamics as an inverse problem. In Mathematical technology of networks, pages 39–48. Springer, 2015. 68

[280] Mark EJ Newman. Network structure from rich but noisy data. Nature Physics, 14(6):542–545, 2018. 68

[281] Roger Guimerà and Marta Sales-Pardo. Missing and spurious interactions and the reconstruction of complex networks. Proceedings of the National Academy of Sciences, 106(52):22073–22078, 2009. 68

[282] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: uses and interpretations. Neuroimage, 52(3):1059–1069, 2010. 68

[283] Andrej Aderhold, Dirk Husmeier, Jack J Lennon, Colin M Beale, and V Anne Smith. Hierarchical bayesian models in ecology: reconstructing species interaction networks from non-homogeneous species abundance data. Ecological Informatics, 11:55–64, 2012. 68

[284] YX Rachel Wang and Haiyan Huang. Review on statistical methods for gene network reconstruction using expression data. Journal of theoretical biology, 362:53–61, 2014. 68

[285] Juanjuan Zhang, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, Stefano Merler, Cécile Viboud, Alessandro Vespignani, et al. Changes in contact patterns shape the dynamics of the covid-19 outbreak in china. Science, 2020. 69

[286] Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov

outbreak originating in wuhan, china: a modelling study. The Lancet, 395(10225):689–697, 2020. 69

[287] Henrik Salje, Cécile Tran Kiem, Noémie Lefrancq, Noémie Courtejoie, Paolo Bosetti, Juliette Paireau, Alessio Andronico, Nathanaël Hozé, Jehanne Richet, Claire-Lise Dubost, et al. Estimating the burden of sars-cov-2 in france. Science, 2020. 69

[288] Ana S Gonzalez-Reiche, Matthew M Hernandez, Mitchell J Sullivan, Brianne Ciferri, Hala Alshammary, Ajay Obla, Shelcie Fabre, Giulio Kleiner, Jose Polanco, Zenab Khan, et al. Introductions and early spread of sars-cov-2 in the new york city area. Science, 2020. 69

[289] Kimberly A Prather, Chia C Wang, and Robert T Schooley. Reducing transmission of sars-cov-2. Science, 2020. 69

[290] Nick Warren Ruktanonchai, JR Floyd, Shengjie Lai, Corrine Warren Ruktanonchai, Adam Sadilek, Pedro Rente-Lourenco, Xue Ben, Alessandra Carioli, Joshua Gwinn, JE Steele, et al. Assessing the impact of coordinated covid-19 exit strategies across europe. Science, 2020. 69

[291] Lucy C Okell, Robert Verity, Oliver J Watson, Swapnil Mishra, Patrick Walker, Charlie Whittaker, Aris Katzourakis, Christl A Donnelly, Steven Riley, Azra C Ghani, et al. Have deaths from covid-19 in europe plateaued due to herd immunity? Lancet (London, England), 2020. 69

[292] C Jessica E Metcalf, Dylan H Morris, and Sang Woo Park. Mathematical models to guide pandemic response. Science, 369(6502):368–369, 2020. 69

[293] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. Science, 368(6489):395–400, 2020. 69

[294] Jonas Dehning, Johannes Zierenberg, F Paul Spitzner, Michael Wibral, Joao Pinheiro Neto, Michael Wilczek, and Viola Priesemann. Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. Science, 2020. 69

[295] Nicholas G Davies, Adam J Kucharski, Rosalind M Eggo, Amy Gimma, W John Edmunds, Thibaut Jombart, Kathleen O'Reilly, Akira Endo, Joel Hellewell, Emily S Nightingale, et al. Effects of non-pharmaceutical interventions on covid-19 cases, deaths, and demand for hospital services in the uk: a modelling study. The Lancet Public Health, 2020. 69

[296] Johannes Haushofer and C Jessica E Metcalf. Which interventions work best in a pandemic? Science, 368(6495):1063–1065, 2020. 69

[297] Aurelio Tobías, Cristina Carnerero, Cristina Reche, Jordi Massagué, Marta Via, María Cruz Minguillón, Andrés Alastuey, and Xavier Querol. Changes in air quality during the lockdown in barcelona (spain) one month into the sars-cov-2 epidemic. Science of the Total Environment, page 138540, 2020. 69

[298] Aiymgul Kerimray, Nassiba Baimatova, Olga P Ibragimova, Bauyrzhan Bukenov, Bulat Kenessov, Pavel Plotitsyn, and Ferhat Karaca. Assessing air quality changes in large cities during covid-19 lockdowns: The impacts of traffic-free urban conditions in almaty, kazakhstan. Science of the Total Environment, page 139179, 2020. 69

[299] Rui Bao and Achen Zhang. Does lockdown reduce air pollution? evidence from 44 cities in northern china. Science of the Total Environment, page 139052, 2020. 69

[300] Zander S. Venter, Kristin Aunan, Sourangsu Chowdhury, and Jos Lelieveld. Covid-19 lockdowns cause global air pollution declines. Proceedings of the National Academy of Sciences, 117(32):18984–18990, 2020. 69

[301] Guillaume P. Chossière, Haofeng Xu, Yash Dixit, Stewart Isaacs, Sebastian D. Eastham, Florian Allroggen, Raymond L. Speth, and Steven R. H. Barrett. Air pollution impacts of covid-19–related containment measures. Science Advances, 7(21), 2021. 69

[302] Pengfei Wang, Kaiyu Chen, Shengqiang Zhu, Peng Wang, and Hongliang Zhang. Severe air pollution events not avoided by reduced anthropogenic activities during covid-19 outbreak. Resources, Conservation and Recycling, 158:104814, 2020. 69

[303] J Yletyinen, GLW Perry, P Stahlmann-Brown, R Pech, and JM Tylianakis. Multiple social network influences can generate unexpected environmental outcomes. Scientific reports, 11(1):1–14, 2021. 69, 76

[304] Martin Heimann and Markus Reichstein. Terrestrial ecosystem carbon dynamics and climate feedbacks. Nature, 451(7176):289–292, 2008. 69

[305] Lombardia ARPA Lombardia-Regione. Rapporto sulla qualità dell'aria di milano e provincia: Anno 2002, 2002. 70

[306] Google LLC "Google COVID-19 Community Mobility Reports". https://www.google.com/covid19/mobility/Accessed:<07-29-2020>, 2020. 71, 79

[307] Retrieved on July 2020 from Apple. https://www.apple.com/covid19/mobility, 2020. 71, 79

[308] Retrieved on July 2020 from Facebook. https://data.humdata.org/dataset/movement-range-maps, 2020. 71, 79

[309] Kenneth O McGraw and Seok P Wong. A common language effect size statistic. Psychological bulletin, 111(2):361, 1992. 71

[310] Robert Coe. It's the effect size, stupid: What effect size is and why it is important, 2002. 71

[311] Bruce L Brown and Suzanne B Hendrix. Partial correlation coefficients. Encyclopedia of statistics in behavioral science, 2005. 71

[312] Michael A Fligner and Timothy J Killeen. Distribution-free two-sample tests for scale. Journal of the American Statistical Association, 71(353):210–213, 1976. 73, 80

[313] Sheldon M. Ross. Introduction to Probability Models. Academic Press, San Diego, CA, USA, sixth edition, 1997. 73

[314] Norman Cliff. Dominance statistics: Ordinal analyses to answer ordinal questions. Psychological bulletin, 114(3):494, 1993. 73, 80

[315] Kenneth O McGraw and Seok P Wong. A common language effect size statistic. Psychological bulletin, 111(2):361, 1992. 73, 80

[316] Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. Science Advances, 5(11):eaau4996, 2019. 74

[317] Jean-Gabriel Young, George T Cantwell, and MEJ Newman. Bayesian inference of network structure from unreliable data. Journal of Complex Networks, 8(6):cnaa046, 2020. 74

[318] Austan Goolsbee and Chad Syverson. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. Journal of Public Economics, 193:104311, 2021. 78

[319] Retrieved on July 2020 from ESA. https://scihub.copernicus.eu/, 2020. 79

[320] Retrieved on July 2020 from Copernicus Climate Change Service. https://cds.climate.copernicus.eu/, 2020. 79

[321] Retrieved on July 2020 from Terrna. https://www.terna.it/en/electric-system/transparency-report, 2020. 79

[322] Gemma Lancaster, Dmytro Iatsenko, Aleksandra Pidde, Valentina Ticcinelli, and Aneta Stefanovska. Surrogate data for hypothesis testing of physical systems. Physics Reports, 748:1–60, 2018. 81

[323] Qingru Sun, Xiangyun Gao, Shaobo Wen, Zhihua Chen, and Xiaoqing Hao. The transmission of fluctuation among price indices based on granger causality network. Physica A: Statistical Mechanics and its Applications, 506:36–49, 2018. 81

[324] Petros Damos. Using multivariate cross correlations, granger causality and graphical models to quantify spatiotemporal synchronization and causality between pest populations. BMC ecology, 16(1):1–17, 2016. 81

[325] James Douglas Hamilton. Time series analysis. Princeton university press, 1994. 81

[326] Filipi N Silva, Didier A Vega-Oliveros, Xiaoran Yan, Alessandro Flammini, Filippo Menczer, Filippo Radicchi, Ben Kravitz, and Santo Fortunato. Detecting climate teleconnections with granger causality. arXiv preprint arXiv:2012.03848, 2020. 81

[327] Allan W Gregory and Michael R Veall. Formulating wald tests of nonlinear restrictions. Econometrica: Journal of the Econometric Society, pages 1465–1468, 1985. 82

[328] Kay H Brodersen, Fabian Gallusser, Jim Koehler, Nicolas Remy, Steven L Scott, et al. CausalImpact 1.2.1. http://google.github.io/CausalImpact/, 2015. 83

[329] Graeme Burrows, Carlos Diuk, Alex Dow, Ismail Onur Filiz, Andreas Gros, Molly Cohn Jackman, Chang Lang, Paige Maas, Waqar Malik, Winter Mason, Chaya Nayak, Drishtie Patel, and Varun Sharma. Facebook disaster maps: Methodology. https://research.facebook.com/blog/2017/6/facebook-disaster-maps-methodology/. Accessed: 2022-03-18. 86

[330] Paige Maas, Shankar Iyer, Andreas Gros, Wonhee Park, Laura McGorman, Chaya Nayak, and P Alex Dow. Facebook disaster maps: Aggregate insights for crisis response & recovery. In KDD, volume 19, page 3173, 2019. 86

[331] Kenneth R. Knapp, Michael C. Kruk, David H. Levinson, Howard J. Diamond, and Charles J. Neumann. The international best track archive for climate stewardship (ibtracs): Unifying tropical cyclone data. Bulletin of the American Meteorological Society, 91(3):363 – 376, 2010. 86

[332] Kenneth R. Knapp, Michael C. Kruk, David H. Levinson, Howard J. Diamond, and Charles J. Neumann. Project, version 4.0. Bulletin of the American Meteorological Society, 2010. 86

[333] Matti Kummu, Maija Taka, and Joseph HA Guillaume. Gridded global datasets for gross domestic product and human development index over 1990–2015. Scientific data, 5(1):1–15, 2018. Accessed: 2020-12-14. 87

[334] Michael Bailey, Ruiqing Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Measuring social connectedness. Technical report, National Bureau of Economic Research, 2017. 87

[335] Facebook social connectedness index. `https://data.humdata.org/dataset/social-connectedness-index?fbclid=IwAR1Egmn1M7WBys362gofAdq19_LpZA8TOM_D5aml2iUD5mqxvxZ3QUgy2oI`. Accessed: 2020-12-17. 87

[336] Ian Christopher Naungayan Rocha, Ana Carla dos Santos Costa, Zarmina Islam, Shubhika Jain, Samarth Goyal, Parvathy Mohanan, Mohammad Yasir Essar, and Shoaib Ahmad. Typhoons during the covid-19 pandemic in the philippines: impact of a double crises on mental health. Disaster Medicine and Public Health Preparedness, pages 1–4, 2021. 89, 91

[337] Bikash Ranjan Mishra and Pabitra Kumar Jena. Bilateral fdi flows in four major asian economies: a gravity model analysis. Journal of Economic Studies, 2019. 89

[338] James Gareth, Witten Daniela, Hastie Trevor, and Tibshirani Robert. An introduction to statistical learning: with applications in R. Spinger, 2013. 89

[339] J. Gutierrez and H. Beech. Typhoon goni leaves philippines after only grazing manila. New York Times, 2020, Accessed MArch 7, 2021. 91

[340] Mohd Fadzil Abdul Rashid and Ishak Ab Ghani. The importance of internal migration in urban planning process: The case study of klang valley region. International Journal of Social Planning and Development, 2011. 94

[341] Matthias Bernt. Migration and strategic urban planning: The case of leipzig. disP-The Planning Review, 55(3):56–66, 2019. 94

[342] Reazul Ahsan, Sadasivam Karuppannan, and Jon Kellett. Climate migration and urban planning system: A study of bangladesh. Environmental Justice, 4(3):163–170, 2011. 94

[343] Zeinab Ebrahimpour, Wanggen Wan, José Luis Velázquez García, Ofelia Cervantes, and Li Hou. Analyzing social-geographic human mobility patterns using large-scale social media data. ISPRS International Journal of Geo-Information, 9(2):125, 2020. 94

[344] 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. Physical Review X, 10(4):041055, 2020. 94

[345] 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):e1003716, 2014. 94

[346] Giulia Pullano, Eugenio Valdano, Nicola Scarpa, Stefania Rubrichi, and Vittoria Colizza. Population mobility reductions during covid-19 epidemic in france under lockdown. MedRxiv, 29:2020, 2020. 94

[347] Antoine Pécoud and Paul De Guchteneire. International migration, border controls and human rights: Assessing the relevance of a right to mobility. Journal of Borderlands Studies, 21(1):69–86, 2006. 94

[348] Heaven Crawley and Brad K Blitz. Common agenda or europe's agenda? international protection, human rights and migration from the horn of africa. Journal of Ethnic and Migration Studies, 45(12):2258–2274, 2019. 94

[349] Cathryn Costello and Itamar Mann. Border justice: Migration and accountability for human rights violations. German Law Journal, 21(3):311–334, 2020. 94

[350] Jennifer Hyndman. The geopolitics of migration and mobility. Geopolitics, 17(2):243–255, 2012. 94

[351] Paul A Kramer. The geopolitics of mobility: Immigration policy and american global power in the long twentieth century. The American Historical Review, 123(2):393–438, 2018. 94

[352] Tanya Basok and Guillermo Candiz. Containing mobile citizenship: changing geopolitics and its impact on solidarity activism in mexico. Citizenship Studies, 24(4):474–492, 2020. 94

[353] Tommaso Agasisti and Antonio Dal Bianco. Determinants of college student migration in italy: Empirical evidence from a gravity approach. Available at SSRN 1063481, 2007. 94

[354] Hugo Barbosa, Marc Barthelemy, 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. 94, 96

[355] Aba Schwartz. Interpreting the effect of distance on migration. Journal of political economy, 81(5):1153–1169, 1973. 94

[356] Filippo Simini, Amos Maritan, and Zoltán Néda. Human mobility in a continuum approach. PloS one, 8(3):e60069, 2013. 95

[357] Chaogui Kang, Yu Liu, Diansheng Guo, and Kun Qin. A generalized radiation model for human mobility: spatial scale, searching direction and trip constraint. PloS one, 10(11):e0143500, 2015. 95, 120

[358] Christian M Alis, Erika Fille Legara, and Christopher Monterola. Generalized radiation model for human migration. Scientific Reports, 2021. 95, 96, 98

[359] Caleb Robinson and Bistra Dilkina. A machine learning approach to modeling human migration. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, pages 1–8, 2018. 98

[360] Wally R Gilks, Nicky G Best, and KKC Tan. Adaptive rejection metropolis sampling within gibbs sampling. Journal of the Royal Statistical Society: Series C (Applied Statistics), 44(4):455–472, 1995. 100

[361] RM Soneira and PJE Peebles. A computer model universe-simulation of the nature of the galaxy distribution in the lick catalog. The Astronomical Journal, 83:845–860, 1978. 101, 121

[362] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. SIAM review, 51(4):661–703, 2009. 101

[363] Simona Fratianni and Fiorella Acquaotta. The climate of italy. In Landscapes and landforms of Italy, pages 29–38. Springer, 2017. 107

[364] L Giordano, F Giordano, S Grauso, M Iannetta, M Sciortino, L Rossi, and G Bonati. Identification of areas sensitive to desertification in sicily region. ENEA, Centro Ricerche Casaccia, Via Anguillarese, 301:00060, 2002. 107

[365] Giuseppina Crescimanno and Kenneth B Marcum. Irrigation, salinization and desertification. Evolution of cropping systems as affected by climate change. Aracne Ed., Roma, Italy, 2009. 107

[366] Roger Guimerà, Ignasi Reichardt, Antoni Aguilar-Mogas, Francesco A Massucci, Manuel Miranda, Jordi Pallarès, and Marta Sales-Pardo. A bayesian machine scientist to aid in the solution of challenging scientific problems. Science advances, 6(5):eaav6971, 2020. 120

[367] Inho Hong, Woo-Sung Jung, and Hang-Hyun Jo. Gravity model explained by the radiation model on a population landscape. PloS one, 14(6):e0218028, 2019. 121

[368] Rien Van de Weygaert and Willem Schaap. The soneira-peebles model, 2007. , 121, 122