IMT School for Advanced Studies, Lucca
Lucca, Italy

Structure, Vulnerability, and Dynamics in Socio-Economic and Ecological Networks

PhD Program in Systems Science

Track in Complex Systems and Networks

XXXVIII Cycle

By

Emanuele Calò

2026

The dissertation of Emanuele Calò is approved.

PhD Program Coordinator: Alberto Bemporad, IMT School for Advanced Studies Lucca

Advisor: Prof. Angelo Facchini, IMT School for Advanced Studies Lucca

The dissertation of Emanuele Calò has been reviewed by:

Giulio Cimini, Physics Department University of Rome Tor Vergata

Antonio Scala, CNR Institute for Complex Systems

IMT School for Advanced Studies Lucca
2026

私の母へ

Contents

List of Figures

List of Tables

Acknowledgements

The work presented in the chapters of this thesis is the result of collaborative research and is based on the following co-authored articles:

- V. D. P. Servedio, A. Bellina, E. Calò & G. De Marzo, "Fitness centrality: a non-linear centrality measure for complex networks," *Journal of Physics: Complexity*, IOP Publishing, vol. 6, num. 1, pp. 015002, 2025.
- E. Calò, G. De Marzo & V. D. P. Servedio, "Species vulnerability and ecosystem fragility: A dual perspective in food webs," *Chaos, Solitons & Fractals*, Elsevier, vol. 199, pp. 116741, 2025.
- E. Calò & A. Facchini, "Multilayer Network Analysis of European Regional Flows," *Entropy*, MDPI, vol. 27, article num. 978, 2025.

I acknowledge the use of LLMs for text refinement.

Qualche ringraziamento per le tante persone che sono state parte di questo lungo ed incredibilmente quasi già finito percorso di circa tre anni e mezzo.

Devo e voglio assolutamente iniziare ringraziando il mio Angelo per il costante supporto e la sempre presente disponibilità: Forza Lupa!

Grazie a NETWORKS, la migliore unità di ricerca di IMT, che mi ha sempre fatto sentire a casa.

Grazie al CSH, a Vito, a Giordano e a tutti i fosterer del CREF.

Grazie al 38esimo ciclo, per tutti i bei momenti, specialmente nel primo anno, che è stato anche quello in cui ne avevo più bisogno.

Grazie al gruppo Napoli, che è stato parte integrante della mia quotidianità per tutto il secondo anno.

Grazie agli amici che ho ritrovato e a quelli che ho incontrato a Vienna, oramai la mia seconda casa.

Grazie a Susanne per il percorso che stiamo facendo insieme.

Grazie a Gabriele e a Michele, che hanno allenato il mio corpo sulla terra, nell'aria e nell'acqua (adesso punto su quest'ultima).

Grazie al dojo che ho ritrovato e che mi permette di seguire la via che risuona con me.

Grazie agli amici del cuore, sparsi un po' in giro, che sono i miei punti di riferimento.

Grazie alla mia famiglia, in particolare a mia mamma, inesauribile fonte d'amore e di ammirazione: spero che, da grande, sarò forte come te.

Grazie a mia nonna, che non c'è più, ma che sarebbe (ed è) fiera di me.

Grazie a me, che, una volta raggiunto un obiettivo, continuo ad andare avanti verso il prossimo.

今度は今度。今は今。

Vita

April 5, 1998	Born, Rome, Italy
2020	Bachelor's Degree in Physics Final mark: 110/110 cum laude Università degli 'Studi Roma Tre', Rome, Italy
2022	Master's Degree in Physics of Complex Systems Final mark: 110/110 cum laude Politecnico di Torino, Turin, Italy
2022	Master's Degree in Physique Fondamentale et Applications : Systèmes Complexes Final mark: Très Bien Sorbonne Université, Paris, France
2025	Intern Complexity Science Hub Vienna, Austria
2025	Visiting Researcher Copernicus Institute of Sustainable Development at Utrecht University Utrecht, Netherlands

Publications

1. V. D. P. Servedio, A. Bellina, E. Calò & G. De Marzo, "Fitness centrality: a non-linear centrality measure for complex networks," *Journal of Physics: Complexity*, IOP Publishing, vol. 6, num. 1, pp. 015002, 2025.
2. E. Calò, G. De Marzo & V. D. P. Servedio, "Species vulnerability and ecosystem fragility: A dual perspective in food webs," *Chaos, Solitons & Fractals*, Elsevier, vol. 199, pp. 116741, 2025.
3. E. Calò & A. Facchini, "Multilayer Network Analysis of European Regional Flows," *Entropy*, MDPI, vol. 27, article num. 978, 2025.

Presentations

1. E. Calò, "Measuring Fitness and Complexity of Species in Food Webs," at *Economic and Financial Networks satellite of NetSci*, Vienna, Austria, 2023.
2. E. Calò, "Multidimensional Territorial Attractiveness: an Application to European Flows," at *PhD and Early Researchers Workshop*, Vienna, Austria, 2024.
3. E. Calò, "Multidimensional Territorial Attractiveness: an Application to European Flows," at *63rd ERSA (European Regional Science Association) congress*, Terceira Island, Portugal, 2024.
4. E. Calò, "Multidimensional Territorial Attractiveness: an Application to European Flows," at *45th AISRe (Associazione Italiana di Scienze Regionali) conference*, Turin, Italy, 2024.
5. E. Calò, "Measuring Fitness and Vulnerability of Species in Food Webs," at *Ecological networks: a window on our changing planet workshop*, Lucca, Italy, 2024.
6. E. Calò, "Multilayer Network Analysis of European Regional Flows," at *Pathways to Future Regional Development workshop (WIFO)*, Vienna, Austria, 2025.
7. E. Calò, "Species Vulnerability and Ecosystem Fragility: A Dual Perspective in Food Webs," at *Complexity Science Hub*, Vienna, Austria, 2025.
8. E. Calò, "Multilayer Network Analysis of European Regional Flows," at *Complexity in Economics & Finance workshop (StatPhys29 satellite)*, Rome, Italy, 2025.
9. E. Calò, "Species Vulnerability and Ecosystem Fragility: A Dual Perspective in Food Webs," at *CCS (Conference on Complex Systems)*, Siena, Italy, 2025.

Abstract

The applicability of state-of-the-art network algorithms is often constrained by specific network architectures, limiting our understanding of complex systems. This thesis addresses two such challenges. First, we tackle the critical limitation of the influential Economic Complexity Index (ECI) and Economic Fitness and Complexity (EFC) algorithms, which are restricted to purely bipartite networks. The method involves introducing a formal generalization of these algorithms, extending their framework to monopartite networks. The principal finding is a novel centrality measure, termed fitness centrality, which identifies "crucial" nodes that serve as essential hubs for a network's dependent members. To demonstrate its utility, this framework is applied to ecological food webs, successfully characterizing species by both their systemic importance and their vulnerability to extinction. Second, we address the inadequacy of single-layer analyses for modeling multifaceted socio-economic interactions. Our procedure is to construct and analyze a multilayer network of European regional flows, integrating data on investment, migration, and other interactions. The analysis reveals hidden interdependencies, identifies versatile regional hubs invisible to single-layer perspectives, and uncovers functionally integrated communities. This approach provides a more holistic and structurally accurate understanding of the European economy. Together, these contributions advance network science by developing more versatile analytical tools and applying them to provide deeper insights into both ecological stability and regional economics.

Chapter 1

Introduction

The study of complex systems has been fundamentally transformed by the rise of network science (Albert and Barabási, 2002; M. Newman, 2018). This field represents a paradigm shift, moving the focus from the analysis of individual components to the intricate web of their interactions. Its power lies in revealing emergent properties (Barabási and Albert, 1999)–system-level behaviors that are not present in the components themselves but arise from their interconnected structure. This network-centric perspective was catapulted into prominence by pioneering algorithms like Google's PageRank (Brin and Page, 1998) and Kleinberg's Hubs and Authorities concept (Kleinberg, 1999). The revolutionary insight shared by these methods was to determine the importance of a node not by its intrinsic properties, but by the quality and quantity of its connections. This principle–that relevance is a relational property defined by the graph structure–proved to be remarkably versatile. The success of this approach stems from the fact that most real-world complex systems exhibit significant and interpretable non-random structures (Caldarelli, 2007; Cimini et al., 2019). Consequently, network-based methods have been widely adopted across diverse domains, from biological and spatial networks to economics and social media (Girvan and M. E. Newman, 2002; Barthélemy, 2011; Schweitzer et al., 2009). They have enabled the identification of key species in food webs (Allesina and

Pascual, 2009), systemically important financial institutions (Battiston et al., 2012), and influential social media users (Weng et al., 2010), demonstrating the profound impact of network science in decoding the complexity of our world.

However, the universal principles of network science must often be adapted to the specific architectures of real-world systems. A particularly common and important structure is the bipartite network, which consists of two distinct sets of nodes where links only exist between the sets, not within them. This two-sided structure is fundamental to diverse systems, such as mutualistic relationships between plants and pollinators or interactions between countries and the products they export. Standard ranking algorithms are ill-suited for such systems, as they fail to account for the different nature of the two node classes. This challenge led to the development of specialized algorithms, most notably the Economic Complexity Index (ECI) (César A. Hidalgo and Hausmann, 2009) and the Economic Fitness and Complexity (EFC) (Cristelli et al., 2013; Tacchella, Cristelli, et al., 2012). Both ECI and EFC were introduced to analyze the bipartite country-product network, where a country is linked to the products it exports. Conceived to determine which economies are more industrially advanced and which goods are more sophisticated, these methods provide key insights into the complex interactions that drive economic development. Their success has been profound, leading to their adoption in a wide range of fields, from analyzing urban systems to ranking chess players, proving their value far beyond their original economic context (Domínguez-García and Muñoz, 2015; De Marzo and Vito D. P. Servedio, 2023; Straccamore et al., 2023; Aufiero et al., 2024; César A Hidalgo, 2021). However, the applicability of these powerful algorithms is constrained by a critical limitation: they are defined exclusively for purely bipartite networks. This represents a significant barrier, as many real-world systems are nearly, but not perfectly, bipartite. For example, in a food web, most species function as either prey or predators, but the structure is broken down by species that are both prey and predators. The presence of even a single such link renders ECI and EFC unusable. Previous studies have shown that the GENEPY index (Scia-

rra et al., 2020), a hybrid of ECI and a linearized EFC, can be computed on non-bipartite networks (Costantini et al., 2022). This indicates that both measures have the potential to be significant in a broad range of networks, thereby calling for a more detailed investigation. This thesis directly confronts this gap. Chapter 2 introduces a generalization of both ECI and EFC, extending their framework to monopartite networks. By recasting the algorithms with the network's adjacency matrix, this work extends their applicability to a much broader class of complex systems. This new methodology opens the door to analyzing previously inaccessible systems and provides a novel framework for assessing the roles of individual nodes, a potential that we explore in the ecological context in Chapter 3.

Just as network structures can be specialized, many real-world systems are too complex to be captured by a single set of interactions. This has led to the development of multilayer networks, a framework for modeling systems with multiple, coexisting types of relationships (Bianconi, 2018). In these structures, the ability to move between layers–like a passenger switching from a metro line to a bus route–can create new pathways and dramatically alter system dynamics, often leading to unexpected emergent behaviors like accelerated diffusion (Gomez et al., 2013). This perspective is crucial to reveal hidden correlations and system-wide vulnerabilities that are invisible when each layer is studied in isolation (Boccaletti et al., 2014). This multilayer approach is particularly powerful for understanding complex socio-economic systems, such as the network of European regions. While previous studies have collected and analyzed individual territorial flows such as mobility (Kang et al., 2020), and, in some cases, employed network science (Provenzano, Hawelka, and Baggio, 2018), such single-layer viewpoints cannot capture the critical interplay between them. For instance, a region's importance in the network may be underestimated if it is not a dominant hub in any single flow but acts as a crucial connector across many. Similarly, communities of regions may not be bound by a single strong tie, but by a combination of moderate flows of capital, people, and knowledge. Chapter 4 addresses this challenge by constructing and analyzing a mul-

tilayer network of European regional flows. This integrated framework allows us to answer questions previously inaccessible. By using multilayer centrality measures, we identify regions that function as versatile hubs across the entire system. Furthermore, by detecting multilayer communities, we uncover cohesive economic and social blocks defined by their multifaceted relationships. This approach provides a more robust and holistic picture of European integration, demonstrating the value of multilayer network science for regional analysis and policy.

The following chapters detail the contributions of this thesis, each building upon the last to address the gaps identified above. Chapter 2 confronts the limitations of ECI and EFC by developing a novel generalization that extends their framework to any monopartite network. The key outcome is a new centrality measure, fitness centrality, which identifies "crucial" nodes that act as essential hubs for a network's most dependent members. We validate its utility by demonstrating its superior performance in network vulnerability analysis. Building on this new framework, Chapter 3 applies these tools to a pressing ecological problem. We develop a bi-dimensional approach to characterize species by their importance index (their role as a carbon source) and their robustness index (their resilience), derived from fitness centrality. This provides a more complete picture of ecosystem stability, simultaneously identifying both key and fragile species to better guide conservation efforts. Finally, Chapter 4 moves from single-layer generalizations to the multifaceted challenge of multilayer systems. We construct and analyze a multilayer network of European regions, where layers represent distinct flows like investment and migration. This integrated approach uncovers hidden interdependencies, revealing the structural drivers of the European regional economy that are invisible to any single-layer analysis.

In summary, this thesis contributes to network science through a connected journey from methodological development to real-world application. It begins by generalizing a class of powerful but restricted algorithms, thereby creating new analytical tools. It then applies these tools to address pressing challenges in ecology, demonstrating their practical utility. Finally, it employs multilayer techniques to reveal the hidden

structure of complex socio-economic systems. Together, these contributions highlight the power of a versatile network science approach to decode the intricate systems that shape our world.

Chapter 2

Fitness centrality: a non-linear centrality measure for complex networks

2.1 Introduction

Although network-based algorithms have proven invaluable for analyzing complex systems, their application is often constrained by specific network architectures. As outlined in the introduction of this thesis, a significant challenge lies in applying powerful bipartite ranking algorithms–namely the Economic Complexity Index (ECI) (César A. Hidalgo and Hausmann, 2009) and the Economic Fitness and Complexity (EFC) algorithm (Cristelli et al., 2013; Tacchella, Cristelli, et al., 2012)–to networks that are not strictly bipartite. Although both were developed to analyze country-product trade data, they operate on distinct principles: ECI relies on a linear iterative process, while EFC employs a non-linear approach that has shown superior performance in capturing system complexity and making economic forecasts (Tacchella, Cristelli, et al., 2013; Tacchella, Mazzilli, and Pietronero, 2018). The strict bipartite

requirement has limited their application, despite evidence from related measures such as the GENEPY index that their principles could be meaningful in a wider context (Sciarra et al., 2020; Costantini et al., 2022).

This chapter directly addresses this limitation by developing a formal generalization of both ECI and EFC for monopartite networks. This is achieved by recasting the algorithms to operate on the network's adjacency matrix and by exploiting the non-homogeneous version of EFC (NHEFC) (Vito Domenico Pietro Servedio et al., 2018), which exhibits fewer convergence problems. In this new, generalized context, we refer to the fitness of a node as its fitness centrality. A primary goal of this chapter is to investigate what properties these measures capture in monopartite networks. Through analysis of the Zachary Karate Club network (Zachary, 1977) and other real-world examples, we verify that in monopartite networks, ECI acts as a community detection algorithm. In contrast, fitness centrality identifies what we call "crucial" nodes: those connected to many low-degree nodes, making them essential hubs for the network's most dependent members. To validate this property, we apply fitness centrality to the problem of network attack vulnerability, which refers to the reduction in network performance resulting from the targeted removal of specific vertices or edges (Barabási and Albert, 1999). This issue was extensively explored in Holme et al. (2002), where various strategies were compared in real-world and artificial networks. We demonstrate that a targeted removal strategy based on fitness centrality is highly effective at fragmenting networks, outperforming standard centrality measures in generating isolated nodes, particularly in offline attack scenarios, where the whole attack strategy must be computed once and for all before initiating the attack. These results show that fitness centrality captures complementary features of network topology.

The work presented here broadens the applicability of two prominent algorithms beyond their original economic and bipartite constraints, establishing them as valuable tools for analyzing a wider variety of complex systems. A non-optimized Julia language implementation of the algorithm is available on GitHub (Vito D. P. Servedio, 2024).

2.2 Materials and Methods

Our objective is to extend to mono-partite networks the formalism of Economic Fitness Complexity (EFC) (Tacchella, Cristelli, et al., 2012), traditionally applied to bipartite networks. At the same time, we shall also generalize the Economic Complexity Index (ECI) (César A. Hidalgo and Hausmann, 2009) to mono-partite networks. Given that EFC and ECI have thus far been utilized exclusively in bipartite networks, we begin by revisiting their formalism in this context using the country-product binary matrix $\mathbf{M}$. In this matrix, the entries M_{cp} are set to 1 if country c significantly exports product p, as determined by the Revealed Comparative Advantage (RCA) (Balassa and Noland, 1989), and 0 otherwise. Our choice of the country-product matrix $\mathbf{M}$ is driven by historical reasons, although any other bipartite graph could serve this purpose in the following.

2.2.1 Economic Complexity Index recap

The algorithm defining ECI is rooted in the method of reflections. Consider N_c as the number of countries and N_p as the number of products. The total number of products exported by each country denoted as k_c, and the number of countries exporting a specific product, denoted as k_p, are determined by the expressions:

$$k_c = \sum_{p=1}^{N_p} M_{cp} \quad \text{and} \quad k_p = \sum_{c=1}^{N_c} M_{cp}. \tag{2.1}$$

These quantities signify the *diversification* of country c and the *ubiquity* of product p, respectively. The iterative map used in the algorithm is as follows:

$$\begin{cases} F_c^{(n)} = \frac{1}{k_c} \sum_{p=1}^{N_p} M_{cp} Q_p^{(n-1)} \\ Q_p^{(n)} = \frac{1}{k_p} \sum_{c=1}^{N_c} M_{cp} F_c^{(n-1)}, \end{cases} \tag{2.2}$$

with initial conditions $F_c^{(0)} = k_c$ and $Q_p^{(0)} = k_p$. The iterative map of Eq. (2.2) has to be stopped after a few iterations since it will converge to a constant vector. The number of iterations is chosen arbitrarily. In

one section of the original paper, 18 iterations were used. The terms $F_c^{(n)}$ and $Q_p^{(n)}$ and their interplay at different n are respectively referred to as the Economic Complexity Index (ECI) and Product Complexity Index (PCI). Sorting countries based on their ECI and products based on their PCI results in their rankings. Countries with a high ECI value are generally wealthier and focus on producing products with a high PCI (Mealy, Farmer, and Teytelboym, 2019). These products, in turn, are generally more technologically advanced. This algorithm, denoted as ECI, captures the interplay between a country's diversification and a product's ubiquity. In A.2 we recall the matrix representation of ECI and its link to the spectral properties of graphs.

2.2.2 Economic Fitness Complexity recap

Like ECI, EFC has been used to rank countries and products based on the country-products binary matrix $\mathbf{M}$. The following non-linear map defines EFC:

$$\begin{cases} \tilde{F}_c^{(n)} = \sum_{p=1}^{N_p} M_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \left(\sum_{c=1}^{N_c} M_{cp} / F_c^{(n-1)} \right)^{-1} \end{cases} \quad \rightarrow \quad \begin{cases} F_c^{(n)} = \tilde{F}_c^{(n)} / \langle \tilde{F}_c^{(n)} \rangle_c \\ Q_p^{(n)} = \tilde{Q}_p^{(n)} / \langle \tilde{Q}_p^{(n)} \rangle_p. \end{cases} \tag{2.3}$$

The initial conditions are set as $\tilde{F}_c^{(0)} = 1$ for all countries and $\tilde{Q}_p^{(0)} = 1$ for all products. The idea behind EFC is that a country's fitness is the sum of the complexity of the products it exports (not the average as in ECI) and that the complexity of a product is primarily influenced by countries with the lowest fitness that export it. If a low-fitness country exports a good, this good has low complexity. In the EFC algorithm, fitness and complexity values are iteratively updated until convergence. Due to potential convergence issues in the original EFC algorithm (Pugliese, Zaccaria, and Pietronero, 2016), we shall use a regularized version called the Non-Homogeneous-EFC (NHEFC) (Vito Domenico Pietro Servedio et al., 2018). In this revised version, the map can be expressed as:

$$\begin{cases} F_c^{(n)} = \delta + \sum_{p=1}^{N_p} M_{cp} / P_p^{(n-1)} \\ P_p^{(n)} = \delta + \sum_{c=1}^{N_c} M_{cp} / F_c^{(n-1)} \end{cases} \tag{2.4}$$

Here, the parameter δ is chosen to be smaller than the typical value of M_{cp} (i.e., $\delta \ll 1$). In this new metric, P_p signifies the "simplicity" of a product p, whose complexity being $Q_p = 1/P_p$. The parameter δ accounts for the intrinsic fitness and simplicity of countries and goods respectively. Contrary to the original version, the non-homogeneous map is not defined up to a multiplicative constant, eliminating the need for the normalization procedure (Vito Domenico Pietro Servedio et al., 2018). All over this work, we will fix $\delta = 0.01$ and consider the map converged as soon as the maximum relative change in fitness following one iteration falls below 1%.

2.2.3 Mono-partite representation

The information relevant for a bipartite network is encapsulated in the rectangular matrix $\mathbf{M}$ with dimensions $N_c \times N_p$. The graph itself is represented by its square adjacency matrix $\mathbf{A}$ with dimensions $(N_c + N_p) \times (N_c + N_p)$. For an undirected graph, $\mathbf{A}$ can be expressed as:

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{M} \\ \mathbf{M}^T & \mathbf{0} \end{pmatrix}. \tag{2.5}$$

The diagonal blocks of $\mathbf{A}$ consist of zeros — the first block with dimensions $N_c \times N_c$ and the second with dimensions $N_p \times N_p$. Given the bipartite nature of the graph, no links exist between countries or between products.
Now, we express both ECI and NHEFC in terms of the graph's adjacency matrix $\mathbf{A}$.

ECI

We introduce the vector $\vec{V}$ with $N_c + N_p$ components, where the first N_c elements represent economic complexity and the last N_p elements represent product complexity. Additionally, we introduce the vector $\vec{k}$, where the first N_c elements denote the diversification of countries and the last N_p components represent the ubiquities of products. The ECI algorithm

can be written as:

$$V_i^{(n)} = \frac{1}{k_i} \sum_{j=1}^{N_c+N_p} A_{ij} V_j^{(n-1)} \quad \text{with} \quad V_i^{(0)} = k_i. \tag{2.6}$$

By introducing the diagonal matrix $\mathbf{K}$ with the elements of vector $\vec{k}$ on its diagonal, we can express ECI in matrix form:

$$\vec{V}^{(n)} = \mathbf{K}^{-1}\mathbf{A}\ \vec{V}^{(n-1)} \quad \text{with} \quad \vec{V}^{(0)} = \vec{k}. \tag{2.7}$$

The matrix $\mathbf{K}^{-1}\mathbf{A}$, denoted as $\mathbf{S}$ hereafter, is a normal matrix, as its rows sum to one. Such matrices, also known as transition matrices in physics, are employed to model random walkers on a network. In essence, equations like the one presented in Eq. (2.7) describe a random walker in its n-th step. The analogy between ECI and a random walk was already pointed out in the original article (César A Hidalgo et al., 2007). The reformulation of ECI in terms of the adjacency matrix of the bipartite graph helps us understand better that ECI can also be considered a spectral community detection algorithm in graphs (Mealy, Farmer, and Teytelboym, 2019; Capocci et al., 2005; V. D. P. Servedio et al., 2005). It might seem that the method of reflections is an unnecessary complication arising from the use of the reduced matrix $\mathbf{M}$ instead of the adjacency matrix $\mathbf{A}$ since all the essential information regarding ECI is already encapsulated in the matrix $\mathbf{S}$. In practice, there are some advantages to working with the square of $\mathbf{S}$. This choice is motivated by the fact that the spectrum of $\mathbf{S}$ is symmetric (if λ is an eigenvalue, then $-\lambda$ is also), and there are bouncing effects when estimating its eigenvectors recursively. Hence, it makes sense to consider the normal matrix $\mathbf{N}_A = \mathbf{S}^2$, which has eigenvalues between zero and one. This approach effectively splits the problem into two independent random walkers – one for countries and the other for products – by creating a random walker moving two steps at a time on a bipartite network. Moreover, squaring $\mathbf{S}$ decreases the ratio between the second-largest and third eigenvalues — λ_3/λ_2 becomes $(\lambda_3/\lambda_2)^2$ — making the second eigenvector of $\mathbf{N}_A$ already a good approximation of ECI (Caldarelli et al., 2012).

NHEFC

Similar to our approach with ECI and PCI, we define a vector $\vec{V}$ with $N_c + N_p$ components. Here, the first N_c components represent countries' fitness, while the remaining N_p components signify the products' simplicity. Adhering to this convention, the symmetric NHEFC map of Eq. (2.4) can be expressed as:

$$V_i^{(n)} = \delta + \sum_{j=1}^{N_c+N_p} A_{ij}/V_j^{(n-1)} \quad \text{with} \quad V_i^{(0)} = 1. \tag{2.8}$$

Alternatively, in a more compact form:

$$\vec{V}^{(n)} = \delta \cdot \vec{1} + \mathbf{A} \cdot (\vec{V}^{(n-1)})^{-1} \quad \text{with} \quad \vec{V}^{(0)} = \vec{1}, \tag{2.9}$$

where $\vec{1}$ represents the vector with all its components set to one, and $(\vec{V})^{-1}$ is the vector obtained by inverting all the components of $\vec{V}$.

With the introduction of the adjacency matrix $\mathbf{A}$ in both the ECI and NHEFC formulations, it is straightforward to generalize both algorithms to the case of any graph, where $\mathbf{A}$ no longer needs to be bipartite. Without bipartiteness in the graph, we no longer differentiate between ECI and PCI, as well as fitness and simplicity. Thus, we consider the vector $\vec{V}$ solely as ECI or fitness in both scenarios.

2.3 Results

2.3.1 Qualitative Analysis

As a first step to understand the properties of the monopartite version of ECI (Eq. (2.7)) and of fitness centrality (Eq. (2.9)), we apply both algorithms to the well-known Zachary Karate Club member connection network. It portrays social interactions within a karate sports club. This network has gained prominence because it delineates the club's participants, who split into two groups following two instructors. Consequently, it has been extensively used to evaluate community detection

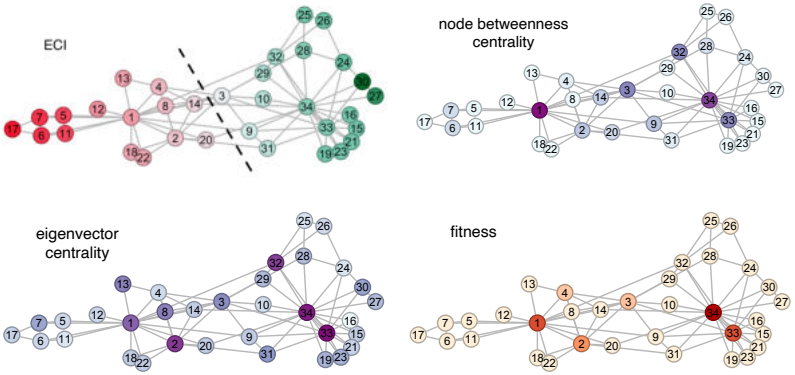

Figure 1: Selected centralities estimated on the Zachary Karate Club network. The nodes appear darker when their respective centrality measures are higher in absolute value. ECI (top-left panel) offers the one-dimensional graph's embedding and identifies its minimum cut. Nodes on either side of the dashed line exhibit opposite-sign ECIs. The minimum cut almost detects the splitting into two separate communities that took place in reality. Only node 3 is wrongly assigned to the component on the figure's right. The node betweenness centrality (top-right) measures the shortest paths going through nodes. This metric is frequently employed to identify nodes whose removal would result in the most significant disruption to the network. Eigenvector centrality (bottom-left) prioritizes nodes with higher degrees and their neighboring nodes. Fitness centrality (bottom-right) is our new proposed centrality measure.

algorithms in graphs (Girvan and M. E. Newman, 2002), under the assumption that this division is an objective measure of the two primary communities. The Zachary Karate Club network is inherently undirected and unweighted. In the upper-left panel of Fig. 1, we depict the Zachary Karate Club network, where nodes' colors are coded according to their ECI value. We observe that ECI effectively identifies the minimum cut of the graph, revealing two distinct communities of nodes. Notably, nodes on opposite sides of the dashed line exhibit ECI values of opposite signs. Interestingly, the "force atlas" method used to plot the network (Bastian, Heymann, and Jacomy, 2009) sets the horizontal position of the nodes in agreement with their respective ECI values. This alignment is not a

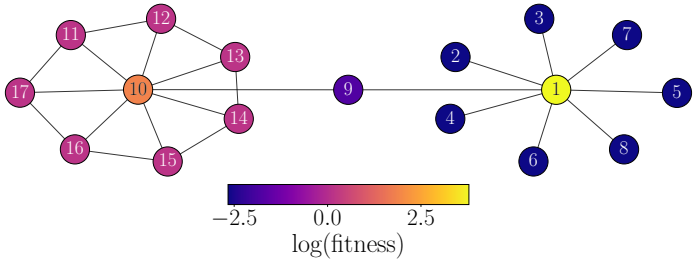

Figure 2: Fitness Centrality in a toy graph. The undirected toy network depicted here consists of a wheel and a star connected by a single node. Both nodes 1 and 10 have a degree of 8. Since nodes with lower degrees generally have lower fitness values, and considering that fitness is calculated based on the sum of the reciprocals of a node's neighbors' fitness values, node 1 acquires a higher fitness than node 10. Higher fitness centrality is depicted with a lighter color.

coincidence, as the force atlas algorithm conceptualizes links within the graph as springs. This analogy invokes the Laplacian matrix, whose first nontrivial eigenvectors delineate the graph's embedding, revealing its underlying community structure. In the upper-right panel of Fig. 1, we employ a color scheme to represent the nodes based on their betweenness centrality within the network. Node betweenness centrality quantifies the shortest paths that traverse a given node. This metric is commonly used to identify nodes with significant traffic flow whose removal could lead to substantial disruptions within the network. Finally, in the lower-left and lower-right panels of Fig. 1, we depict the Zachary Karate Club network with nodes' colors indicating their eigenvector centrality and fitness centrality respectively. Notably, in both scenarios, nodes with high degrees are highlighted prominently. Fitness centrality tends to assign relatively less significance to the neighbors of high-degree nodes than the eigenvector centrality.

To better understand the properties captured by fitness centrality, we computed this measure for the toy graph shown in Fig. 2. This network consists of a star connected to a wheel. Node 1 at the center of the star has higher fitness than node 10 at the center of the wheel because it is

linked to nodes with lower degrees, which in turn have lower fitness values. We recall that a node's fitness centrality is calculated as the sum of the reciprocals of its neighbors' fitness values. The implication is that nodes with high fitness centrality are typically connected to many nodes with a low degree. This is well shown in Fig. 2, where node 1 has a higher fitness centrality than node 10, despite the two nodes being characterized by the same degree.

2.3.2 Correlation with other centrality measures

The qualitative analysis we discussed in the previous section helps us understand the properties captured by the fitness centrality and the features of the monopartite version of the ECI algorithm. However, a more quantitative analysis is needed to capture the correlation of these measures with the other centrality indices. Indeed, it is well known that most centrality measures can present strong correlations, particularly when networks are almost random or show a high spectral gap (Benzi and Klymko, 2015; Oldham et al., 2019). In this paper, we only focus on the fitness centrality and its properties, since the community detection properties of ECI are already well documented.

We start our analysis considering a second toy graph, a linear chain composed of 100 nodes, shown in Fig. 3. As shown in the top panel and the left-bottom one, fitness centrality assigns the largest value to the second and second-last node. This is because these are the only nodes connected with a node with degree one. In the figure's bottom row, we also show a comparison between the fitness centrality and three other centrality measures (degree centrality, eigenvector centrality, and betweenness centrality). We practically observe no correlation between fitness centrality and the rest of the measures, further confirming its ability in capturing different properties of nodes.

As a second step, we focus on a set of well-known and studied real networks: the Zachary Karate Club network we already discussed, the social network of Dolphins, the network of connections between terrorists in a Train bombing attack, the collaboration network of Jazz musi-

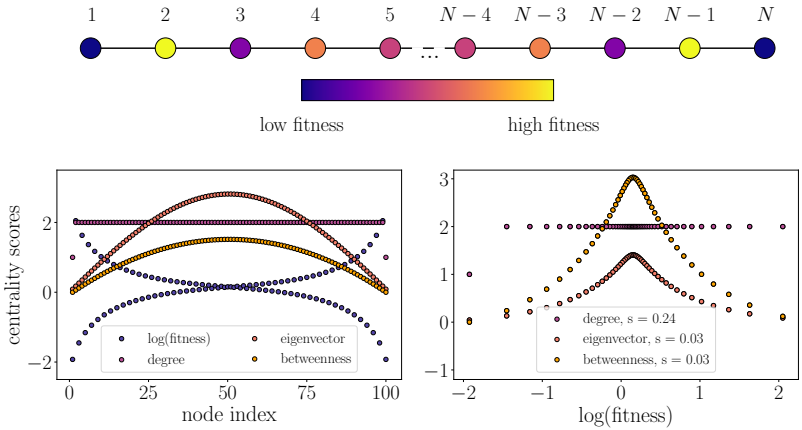

Figure 3: Chain network toy graph. Top: Schematic representation of chain network with nodes colored according to fitness values. **Bottom Left:** Centrality measures values as a function of the node. Note how fitness alternates from low to high values after each node. This is due to the nonlinearity of the fitness centrality algorithm, which assigns to each node the inverse of the score of its neighbors, creating the characteristic alternating pattern. **Bottom Right:** Degree, eigenvector, and betweenness centralities, against fitness centrality. We show the corresponding Spearman rank correlations s in the inset.

cians, and the estimated neural network of the nematode Caenorhabditis Elegans. We considered these five networks as undirected and unweighted. More details can be found in A.1.

We show in Fig. 4 a comparison between fitness centrality and betweenness centrality, eigenvector centrality, and node degree, as estimated in the five considered empirical networks. Each panel in the figure corresponds to a different network. We observe little correlation with the betweenness centrality, with Spearman correlations around $0.5/0.6$. This is not surprising since differently from fitness centrality, this measure is not degree-based but path-based. Correlations around 0.8 are observed with the eigenvector centrality, and the strongest correlations, with values reaching up to 0.9, are observed when comparing the fitness centrality with the degree. In this case, this comes with no surprise, as the same

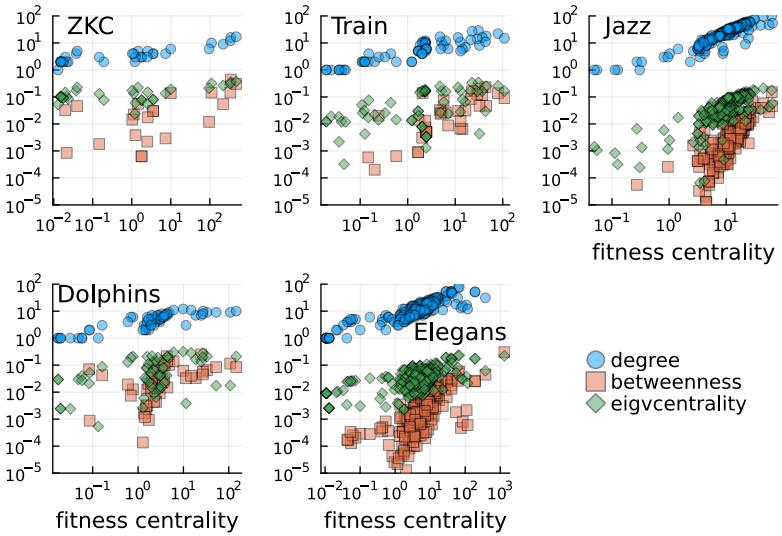

Figure 4: Selected centrality measures vs. fitness centrality in various networks. Each symbol in the plots represents a single node within the network. Across all plots, the horizontal axis represents fitness on a logarithmic scale. We observe a correlation between fitness, node betweenness, eigenvector centrality and node degrees. Fitness does not correlate with ECI (figure in SI), which measures the one-dimensional graph embedding of graphs – a conceptually distinct attribute.

behavior is observed also in bipartite networks. Indeed, fitness centrality coincides, at first order, with the degree. A detailed table (Table 5) reporting all correlations can be found in A.3.1.

We performed a similar analysis also for different synthetic networks, in particular random graphs, Barabási-Albert networks, and small-world networks. Comparing the fitness centrality with the other centrality measures we observe:

- perfect correlation in the random Erdős-Rényi networks. This is expected since in this class of networks there is not much more information than the degree of nodes;

- mild to high correlations in scale-free Barabási-Albert networks.

Also in this case the highest Spearman correlation (0.77) is with the degree;

- high correlation with the degree (Spearman 0.77) but low correlations with the other centrality measures for Watts-Strogatz networks.

The interested reader can refer to the detailed results and plots in A.3.3.

Finally, we tested whether in these networks there is a relation between the fitness centrality and the average nearest-neighbor connectivity (Knn). This is indeed a feature we observed when analyzing the wheel-start toy graph. In particular, we expect a high-fitness node to exhibit a lower Knn. This is the case for the ZKC, Elegans and Dolphin and networks (Spearman correlation -0.88, -0.57, -0.46 respectively), while the other two networks show a less pronounced anti-correlation. Detailed results of this analysis are reported in A.4.

2.3.3 Network Vulnerability

Our analysis of fitness centrality revealed an important insight: this measure effectively identifies nodes connected to low-degree neighbors. However, we also found a correlation between fitness centrality and node degree. This overlap raises an important question about the measure's unique value. To demonstrate that fitness centrality provides meaningful information beyond simple node degree, we need to conduct more detailed research. Our next steps will involve carefully examining what additional network characteristics this metric might capture that cannot be explained by traditional degree measurements. To this extent, we examine network vulnerability as a critical aspect of network analysis. This approach helps us understand how networks respond to disruptions, such as when nodes or connections fail. Our goal is to assess a network's resilience and identify the most effective strategies for targeted interventions that could potentially compromise the network's structure and function. It has been demonstrated that, for synthetic networks as Erdős–Rényi, Barabási–Albert, or Watts-Strogatz, an attack based on betweenness score generally proves more effective in degrading network

performance (Holme et al., 2002). The evaluation of a strategy depends on the metric adopted. Standard choices, as detailed in (Holme et al., 2002), are the size of the largest connected component S_c and the inverse geodesic distance $\langle l^{-1} \rangle$, representing network navigability and functionality. Given the specific feature of fitness centrality, in this study, we also consider another metric, the number of disconnected nodes in the network N_d, which reflects the grade of network disintegration. For example, this can be a relevant measure in the case of supply chains of food webs, where an isolated company or animal can no longer sustain its needs. Specifically, we focus on the fraction of currently remaining connected nodes, denoted $1 - N_d$. Note that we scale the other measures to the range 0 to 1 to facilitate comparison.

In A.5, we provide a comprehensive explanation of the attack procedure. Two attack variants were considered: "initial" (I), where node rankings remain fixed, and "recomputed" (R), where metrics are recalculated after each removal. In Fig. 5, we show the results of network vulnerability to targeted node removal, comparing fitness-based attacks with degree-based and betweenness-based strategies for the I-attack across the same five real-world networks analyzed in Fig. 4. The plots show that the curves can exhibit non-monotonic behavior. This is not an error in the attack strategy, but rather the result of how the metric is normalized at each step. If the removed node was already isolated, the total number of nodes (the denominator of the fraction) decreases, while the number of connected nodes remains constant. This increases the proportion of connected nodes among the remaining population.

We also quantitatively evaluated the performance of fitness, degree, and betweenness centralities using the Area Over the Curve (AOC), which is reported in the plots. Larger AOC values correspond to more successful attacks, corresponding to the network being entirely disrupted at the first node removal. Larger AOC values always reflect higher performance. The highest performance ratings are indicated by bold text. Fitness-based removal consistently outperformed other strategies in terms of N_d. For S_c and $\langle l^{-1} \rangle$, fitness performed better except for the Arena-Jazz network, where betweenness emerges as the optimal strategy. In

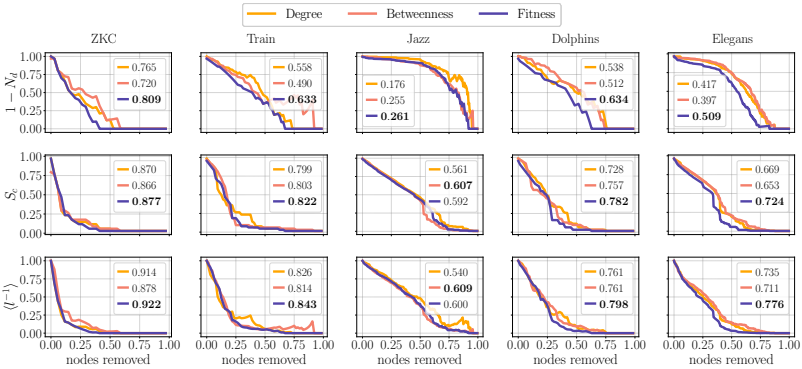

Figure 5: Attack vulnerability of real networks (I-variant). Results of the attack vulnerability analysis with the I-variant (centralities calculated once on the full network) on five different real-world networks. The networks are the same as in Fig. 4. We compare the performance of degree, betweenness, and fitness strategies. Fitness refers to the original metric of Eq. (2.9). Each quantity $(1-N_d)$, S_c, $\langle l^{-1} \rangle$, and the number of nodes removed are normalized to 1 for visualization. Each column shows the result of the analysis for a single network. The fitness strategy generally outperforms the others in terms of $(1 - N_d)$, S_c, and $\langle l^{-1} \rangle$, showing in some cases a significant advantage. In the case of the Arena-Jazz network, betweenness is the best strategy, although the performance is very close to the one obtained via fitness. We denote the values corresponding to the best performance (highest AOC) with a boldface character.

A.5.2, we present an analogous analysis for the R-attack. While we still observe that fitness-based removal consistently outperforms other strategies in terms of $(1 - N_d)$, for S_c and $\langle l^{-1} \rangle$ fitness is generally outperformed by degree and betweenness. Notably, the computationally less demanding I-attack variant using fitness proved highly effective, potentially offering advantages in specific scenarios. For instance, when the network's size makes it impossible to recompute the measure after each removal or in situations like parallel police operations (Cavallaro et al., 2020). In A.5, we also report the results of the same analysis performed on synthetic networks for both I and R-attacks (A.5.1 and A.5.2 respectively), while in A.6 we present the results of the same analysis performed on real weighted networks for the I-attack. Also in these cases,

fitness centrality turns out to be the best metric for N_d and the best metric overall in I-type attacks.

The reason why fitness centrality performs so well in guiding network attacks is easy to understand. As we already stressed, high-fitness nodes, tend to be connected to low-degree nodes. This implies that high-fitness nodes are more likely to disrupt the network, creating many isolated nodes, if removed. This is clear by looking at our toy graph in Fig. 2, removing node 1 detaches seven nodes from the network, whereas removing node 10 detaches the wheel from the network, but its neighbors remain connected in a circle. Analogously, removing the second or second-last node in the chain of Fig. 3, those with the highest fitness, causes the formation of an isolated node.

2.3.4 Generalization to other types of networks

In this paper we focus on unweighted undirected networks, but the fitness centrality algorithm works well also when these assumptions are relaxed and the fitness centrality developed in previous sections can be extended to both weighted and directed networks. In the case of weighted networks, the generalization is straightforward: one uses the weighted adjacency matrix. Two examples of fitness calculations on undirected weighted networks are presented in A.6. In directed networks, the complexity increases due to two distinct types of nodal connections: outgoing and incoming links.

The information that flows through these connections should be treated separately. In real-world scenarios, a node with many outgoing links may exhibit very different properties from a node with many incoming links. For eigenvector centrality, we define left and right centrality, corresponding to the left and right eigenvectors of the network's adjacency matrix $\boldsymbol{A}$. This is equivalent to taking the eigenvector of $\boldsymbol{A}$ and $\boldsymbol{A}^T$, respectively. These two measures provide different descriptions: one related to the inflow of information and the other to the outflow. While left and right eigenvector centralities are positively correlated with in-degree and out-degree, these two types of information may be treated

differently in some cases (Rodgers, Tiňo, and Johnson, 2023). For example, in a food web (Dunne, 2006), a species feeding on many prey (high out-degree) reflects a high rank in the food chain, whereas having many predators (high in-degree) is generally a disadvantage for the species.

Following the same reasoning as the eigenvector centrality, we define from Eq. (2.9) an *out-fitness* $V^{(\text{out})}$ and an *in-fitness* $V^{(\text{in})}$. These are computed by running the algorithm for both $\boldsymbol{A}$ and $\boldsymbol{A}^T$. While it is possible to generalize Eq. (2.9) and calculate two separate independent fitness values – one for the adjacency matrix and another for its transpose – it is often more insightful to combine these two quantities. In this case, the solution at the fixed point is represented by the following system of equations:

$$\begin{cases} V_i^{(\text{out})} = \delta + \sum_{j=1}^{N} A_{ij}/V_j^{(\text{in})} \\ V_i^{(\text{in})} = \delta + \sum_{j=1}^{N} A_{ji}/V_j^{(\text{out})} \end{cases} \tag{2.10}$$

where N is the number of nodes in the network. These two quantities measure the same network features analyzed in the previous sections, but one refers to outgoing connections and the other to incoming connections. This allows us to disentangle the two different flows of information in a directed network.

The application of this algorithm is particularly interesting in the case of food webs and economic supply chains. In Chapter 3, we will deal with the specific case of food webs. A similar argument applies to the supply chain networks of firms. Here, $V^{(\text{out})}$ can be interpreted as related to the firm's *downstreamness*, and $V^{(\text{in})}$ to its *upstreamness*. These two economic concepts measure the effect in terms of down- and up-propagation of shocks due to the failure of a specific firm (Miller and Temurshoev, 2017). Thus, combining $V^{(\text{out})}$ and $V^{(\text{in})}$ could provide an interpretable way to assess the economic systemic risk of firms in the supply chain (Diem et al., 2022).

While Eq. 2.10 couples the two directions, a standalone version—where the index is computed from the adjacency matrix, without feedback from its transpose—can also be defined. For example, this can be done using only $V^{(\text{out})}$. In this configuration, a node achieves high fitness by pointing to neighbors with low out-fitness. This provides a valuable lens for

identifying nodes that dominate the system's terminal periphery. In the domain of logistics, for instance, high out-fitness nodes would highlight last-mile distribution centers that transition goods to final consumers. In this context, the standalone metric identifies the hubs that direct the network's flow to its final destinations.

2.4 Discussion

Estimating the latent capabilities of actors within a network poses considerable challenges. These challenges become even more daunting when we rely solely on the network's topological structure, with limited additional information available. The Economic Complexity Index (ECI) and the Economic Fitness Complexity (EFC) have proven instrumental in addressing this gap. They have been applied to bipartite networks of countries and exported products to indirectly infer countries' inner capabilities. Remarkably, despite economists typically relying on many indicators, the scarce information on exported products can yield impressive results in economic impact forecasting when approached using the ECI and EFC algorithms. Before this study, the applicability of ECI and EFC was limited to bipartite networks and only a few studies generalized similar algorithms to other types of networks (Costantini et al., 2022). We closed this gap by showing how both algorithms can be applied to any graph and utilized the well-known Zachary Karate Club network and other four additional networks, for thorough testing. The resultant generalized algorithms are outlined in Eqs. (2.7) and (2.9). In both instances, we observe a loss in interpreting the Product Complexity Index and Economic Complexity of products, respectively, as we now only require one vector. The dual sets of quantities - ECI and PCI and Economic Fitness and Economic Complexity - were artifacts stemming from the system's bipartite nature.

In the case of mono-partite networks, the generalized ECI is closely linked to the graph's community structure, effectively serving as a spectral-based community detection algorithm. It can also be viewed as a one-dimensional graph embedding and provides the graph's minimum cut.

ECI functions as a random walk on graphs, failing to fully capture the significance of high-degree nodes. For instance, compared to eigenvector centrality — which prioritizes high-degree nodes and their neighbors (refer to the lower left panel of Fig. 1) — ECI assigns more importance to node 30 over node 34 and treats node 1 and 34 differently, despite both having a high number of connections. Indeed, nodes 1 and 34 exhibit significantly different rankings according to ECI despite their roles being nearly equivalent in importance. After all, nodes 1 and 34 are known to be responsible for the split. Conversely, fitness centrality assigns greater importance to pivotal nodes 33, 34, and 1. It comes as no surprise that high-degree nodes hold the utmost importance in the network. This observation has led to criticism of EFC, suggesting it does not offer additional valuable information beyond node degree. Fitness centrality is determined by the sum of the inverse fitness values of neighboring nodes, giving node 1 a higher value. This suggests that fitness centrality incorporates information on the degree of nodes and the connectivity of their nearest neighbors.

To clarify this point, we performed an extensive analysis comparing the fitness centrality to other measures, including the degree, in both real and synthetic graphs. Despite a high correlation with degree centrality, our study revealed that fitness centrality can capture complementary properties. Specifically, nodes with higher degrees and lower nearest-neighbor connectivity tend to have higher fitness centrality. In general, we observe a power-law-like relationship between degree and fitness. The power-law exponents relating fitness to degree are reported in Table 5 in A.3.1. They appear to depend on the number of nodes and links, although we do not have a clear explanation for this phenomenon. However, we can observe that in a graph composed of cliques with different numbers of nodes, the degree scales as the square of the fitness. In a clique, all nodes are equivalent and therefore have the same fitness value, which allows us to write a single quadratic equation and solve for the fitness.

As a last analysis, we proved that fitness centrality is an effective tool for assessing network vulnerability. Depending on the metric adopted

and the attack strategy followed, fitness centrality performs the best or similarly to the best-performing metrics. In particular, when centrality measures are calculated once for the entire network, fitness centrality outperforms other measures. This makes it a fast and efficient method for identifying vulnerable nodes in a network. Moreover, fitness centrality systematically outperforms other measures in creating the largest number of isolated nodes. Fitness centrality is therefore a valuable index for identifying "crucial" nodes, that when removed lead to the isolation of many other nodes. These results point out that fitness centrality carries information beyond the degree. While in centralities like Katz or eigenvector, nodes increase their status by connecting to other high-status nodes, fitness centrality operates conversely: a node's importance is inversely proportional to the importance of its neighbors. This also distinguishes it from other nonlinear measures like p-norm centralities, which typically aim to amplify the influence of the strongest neighbors. In contrast, fitness centrality amplifies the influence of nodes that serve the weakest (least-connected) members of the network.

Moreover, our method is computationally less demanding and could benefit large networks where the computation of node betweenness is slow. High-fitness nodes can be estimated accurately after just a few iterations of the map in Eq. (2.9), while longer computational times are needed to achieve convergence for low-fitness nodes, which have negligible importance for network disruption.

Further analysis could explore modifications of the generalized fitness map. For example, the nonlinear term can be adjusted as $(\vec{V})^{-\nu}$, where $0 < \nu < 1$ diminishes the negative significance of low-degree nodes, while $\nu > 1$ amplifies their importance. This would likely allow to tune the correlation with the degree. Some work going in this direction for the specific bipartite networks case of can be found in (Andrea Mazzolini, 2024).

To summarize, our work presents a significant advancement in network analysis by generalizing ECI and EFC to monopartite networks, thus extending their applicability to virtually any graph structure. This generalization bridges a crucial gap in complex network analysis, allow-

ing these powerful algorithms to be applied beyond their original economic context to several systems. Moreover, we demonstrate the fitness centrality algorithm's effectiveness in identifying "crucial" nodes in various network types and highlight its complementarity to existing network measures.

Chapter 3

Species vulnerability and ecosystem fragility: a dual perspective in food webs

3.1 Introduction

The preservation of ecosystems is one of the most pressing challenges in contemporary ecology. A key aspect of this challenge lies in understanding the consequences of species loss and identifying the most vulnerable species. These tasks are increasingly urgent due to escalating anthropogenic pressures, such as climate change and over-exploitation, which severely disrupt ecosystems (Petchey et al., 1999; Harmon, Moran, and Ives, 2009; Blanchard, 2015; Morris et al., 2016; Martínez-Ramos et al., 2016; Bartley et al., 2019; Nagelkerken et al., 2020; Gomes et al., 2024). Species extinctions often lead to cascading effects throughout the ecosystem, triggering secondary extinctions that amplify the initial damage (Dunhill et al., 2024; Keyes, Barner, and Dee, 2024). To mitigate these impacts, it is essential to quantify species' "relevance" and "robustness," guiding conservation policies to prioritize the preservation of fundamental and fragile species.

Ecosystems, like many complex systems, are characterized by non-

linear interactions and intricate structures (Fussmann and Heber, 2002; Pillai, Gonzalez, and Loreau, 2011), where each component contributes to overall functionality. Similar dynamics are observed in other domains, such as interbank loan markets or supply chains, where the removal of a single bank or firm can significantly disrupt economic output (Karimi and Raddant, 2016; Diem et al., 2022; Stangl et al., 2024). In this context, network science has emerged as a powerful tool for studying such systems. By representing species as nodes and their interactions—such as mutualistic, host-parasite, or trophic relationships—as links, networks provide a mathematical framework for analyzing ecological systems (Eklöf et al., 2013; Domínguez-García and Muñoz, 2015). Among these networks, food webs, which depict "who consumes whom" interactions, have been extensively studied to understand ecosystem dynamics (Dunne, Williams, and Martinez, 2002; Garlaschelli, Caldarelli, and Pietronero, 2003; Krause et al., 2003; Otto, Rall, and Brose, 2007; Allesina, Alonso, and Pascual, 2008; Ghoshal and Barabási, 2011; Allesina, Grilli, et al., 2015; Keyes, McLaughlin, et al., 2021). Network-based measures, particularly those derived from PageRank or eigenvector centrality (Page, 1999; Bryan and Leise, 2006; Thurner, Hanel, and Klimek, 2018), have proven effective in quantifying species' relevance within food webs. These metrics identify species whose removal would have the most significant impact on biodiversity (Allesina and Pascual, 2009; McDonald-Madden et al., 2016). This approach is especially useful for large food webs, where exhaustive simulations of all extinction scenarios are computationally infeasible. However, while these methods excel at identifying critical species, they often overlook those that are highly vulnerable. For instance, some species may play a minor role in the food web economy but remain extremely susceptible to extinction due to specialized diets or habitat requirements—pandas being a notable example.

While network science has provided numerous metrics for quantifying species' roles in food webs, including measures of centrality and structural roles (Estrada, 2007; Luczkovich et al., 2003), and the concept of keystone species has been explored through various network approaches (Jordán, Takacs-Santa, and Molnar, 1999), our methodology of-

fers a distinct perspective. Building on the fitness centrality introduced in Chapter 2, our approach explicitly and simultaneously quantifies two complementary dimensions for each species: the importance index (measuring its centrality as a carbon source and its relevance within the food web) and the robustness index (capturing its predatory ability and resilience to extinction). Unlike single-index measures that characterize different aspects of species' network roles (Cirtwill et al., 2018; Rocchi et al., 2017), this bi-dimensional framework provides a direct and separate assessment of these two fundamental ecological facets for each species, building upon the idea that considering both a species' dependence on resources and its impact on its prey is crucial for a comprehensive understanding of its ecological role, as also explored through combinations of centrality indices (Gouveia, Móréh, and Jordán, 2021). This bi-dimensional characterization then allows us to visualize the food web's structure, where each species is represented as a point on this plane. Our findings reveal that highly important species often drive significant extinction cascades when removed, whereas low-robustness index species exhibit limited resilience to such events. This methodology provides a more nuanced understanding of both critical and vulnerable species within ecosystems. By capturing the complex interplay between predation and survival in food webs, our approach offers a valuable, ecologically-informed perspective that could contribute to the broader framework of conservation strategies, potentially enhancing the prioritization of biodiversity protection efforts.

3.2 Materials and Methods

3.2.1 Robustness index and Importance index of Species

Quantifying species' roles and impacts within ecosystems presents a significant methodological challenge in ecological research. Traditional network-based methodologies approach this task by looking at the food web formed by the species. This consists of a directed network, where nodes are the species, while edges signify carbon flow between these species. As

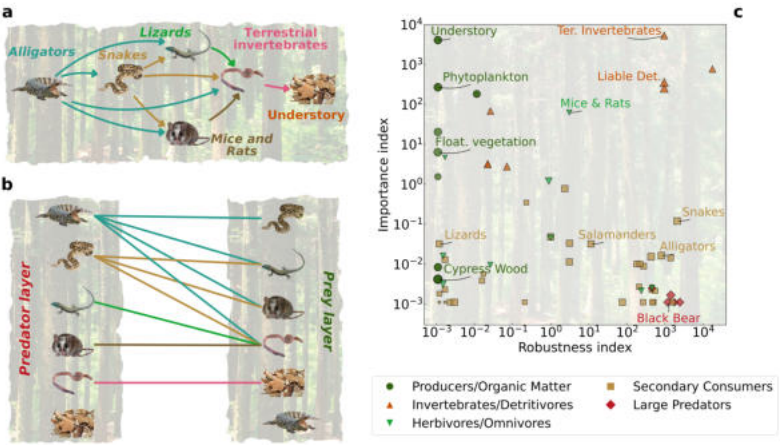

Figure 6: Cypress food web: network representations and robustness index-importance index plane. a, Directed monopartite subnetwork illustrating predator-prey relationships. **b,** Equivalent undirected bipartite representation of the subnetwork. **c,** Robustness index-importance index plane depicting species' ecological roles. Points represent individual species, with colors indicating functional categories and sizes proportional to species biomass. Selected species are labeled for reference. This visualization reveals the dual roles of species as both predators and prey within the ecosystem.

predation represents a major pathway for carbon flow between species, these networks are frequently termed prey-predator networks. We show in Fig. 6a a schematized food web, where we adopt the convention of arrows going from predators to prey. Once a food web has been reconstructed, the standard approach consists in applying a centrality measure to rank the species. For instance, eigenvector centrality has proven very effective in determining which species have a stronger potential to cause chain extinctions if removed from an ecosystem (Allesina and Pascual, 2009). These measures, however, are unable to assess which species are most in danger during these extinction cascades and may not fully capture species' vulnerability or the nuances of direct and indirect interactions, see also mixed trophic impact analysis (Ulanowicz and Puccia, 1990) and loop analysis (Bodini, Giavelli, and Rossi, 1994).

To address this limitation, we introduce a novel algorithmic approach that quantifies species' dual roles through two distinct measures: the importance index and the robustness index. The importance index measures how important in the food web a species is, while the robustness index assesses a species' predatory capabilities and adaptability within the food web. The importance index of a species is high if it serves as prey for multiple species with a low robustness index. On the other hand, the robustness index evaluates a species' predatory efficiency by assessing both the quantity and the importance index of its prey species. Species that prey upon diverse taxa, particularly those of low importance index, demonstrate high predatory capacity and consequently receive high robustness index values. Mathematically, we formulate this framework as a bi-dimensional non-linear map by using Eq. (2.10). Indeed, let $\mathbf{A}$ represent the adjacency matrix of the food web. The element A_{ij} equals one if there is carbon transfer (predation) from species i to j, and zero otherwise. The robustness index (R_i) and importance index (I_i) of species i are computed through the following iterative map:

$$\begin{cases} R_i^{(n+1)} &= \delta + \sum_j A_{ji} \,/\, I_j^{(n)} \\ I_i^{(n+1)} &= \delta + \sum_j A_{ij} \,/\, R_j^{(n)} \end{cases} \quad R_i^{(0)} = I_i^{(0)} = 1 \;\forall i. \tag{3.1}$$

Here, $R_i^{(n)}$ and $I_i^{(n)}$ denote the robustness index and the importance index of the species i at iteration n, respectively. The quantity δ is a regularization term that guarantees convergence. It does not affect the final ranking, as soon as it is way smaller than $\mathbf{A}$'s elements. In this paper, we set it at 10^{-3}. The algorithm iterates until the convergence of both quantities. The first equation tells us that a species has a high robustness index if it can absorb carbon from several sources, in particular from those with a low importance index. Referring to the second equation, species with a low importance index are consumed by few species, which suggests they are harder to prey upon. Conversely, the second equation implies that the importance index of a species is high if it is consumed by several other species, particularly those with a low robustness index. Low-robustness index species can indeed only absorb carbon from a limited set of sources, so the extinction of any of those would put them in seri-

ous danger. For this reason, we identify vulnerability as the inverse of the robustness index. The lower the robustness index of a species, the more vulnerable to food web shocks it is expected to be.

This approach shares notable similarities with the algorithms developed for bipartite networks in the field of Economic Complexity (César A. Hidalgo and Hausmann, 2009; Tacchella, Cristelli, et al., 2012; Vito Domenico Pietro Servedio et al., 2018; Andrea Mazzolini, 2024), specifically with the algorithms of Economic Fitness and Complexity (EFC) (Tacchella, Cristelli, et al., 2012; Vito Domenico Pietro Servedio et al., 2018). While the EFC algorithm has successfully analyzed various bipartite nested networks, including ecological systems (Domínguez-García and Muñoz, 2015), its application to mono-partite directed networks represents a novel extension, as shown in Chapter 2. In ecological contexts, nestedness contributes significantly to ecosystem stability, biodiversity, and resilience, particularly in mutualistic networks where specialist species typically interact with generalist ones (Mariani et al., 2019). This similarity between the two approaches is not surprising, since Fig. 6b illustrates how the directed trophic interactions of a food web can be visually represented using a bipartite layout, which can be useful for specific analytical approaches. Differently from standard bipartite networks, layers contain the same set of species. The first looks at their role as predator, and the second one looks at their role as prey. Links between the two layers connect predators to their prey.

3.2.2 Networks Construction and Preprocessing

Food webs are represented as directed monopartite networks, where edges denote energy flow between species. In our binary model, an edge from node i to node j indicates i preys upon species j, with no weights assigned to the links. Self-loops are included to account for cannibalism within species. We selected six food webs (see B.1 for details) for analysis and modified them to ensure irreducibility and primitivity, following the method described by (Allesina and Pascual, 2009; Allesina and Bodini, 2004; Allesina, Bodini, and Pascual, 2009). This modification process

involved:

1. Adding a root node to represent the external environment.
2. Creating links from primary producers to the root, reflecting the origin of matter in the food web.
3. Establishing links from the root to every node, representing intrinsic matter loss and detritus accumulation, which is partially recycled back into the food web.

The macroscopic properties of these modified networks, including the number of nodes, links, and network density, are detailed in B.2.

3.2.3 Robustness index-Importance index Plane

To categorize the species of the Cypress food web, we used an open-ended prompt with ChatGPT-4o (accessed on 24/10/2024): *'Can you classify the nodes in this food web into 5 categories?'*, followed by the list of species in the Cypress food web. This approach was chosen to allow the model to leverage its internal knowledge of the specific species to identify their functional roles without being biased by pre-defined labels. Despite the lack of specific instructions, the model's classification (see B.7 for the complete species categorizations) naturally emerged around functional trophic guilds, ranging from Producers to Large Predators. The choice of five categories was a deliberate constraint to align the classification with the typical hierarchical resolution of a complex wetland ecosystem. This provided a manageable number of groups for comparative analysis of importance and robustness while maintaining the necessary ecological detail to distinguish between different levels of the food web.

To enable cross-network comparisons, we normalize the robustness index and importance index values for each food web using the following procedure:

1. Compute the robustness index and importance index for each species.
2. Take the logarithm of these values.

3. Scale the log-transformed values to the [0,1] interval using min-max normalization:

$$x_{\text{normalized}} = \frac{\log(x) - \min(\log(x))}{\max(\log(x)) - \min(\log(x))}.$$

This approach preserves relative differences while allowing for consistent comparison across different food webs.

3.2.4 Extinction Areas

We implement the following species removal procedure to assess ecosystem robustness. Species are progressively removed based on rankings from specified algorithms, maintaining fixed rankings throughout the process. Cascading extinctions are monitored, with a species considered extinct when left without outgoing links or with only a self-loop. The procedure continues until all species go extinct. An extinction curve is generated by plotting the fraction of extinct species against the number of removed species. The extinction area, calculated as the integral of this curve, measures the ecosystem vulnerability. This area equals 1 when all species go extinct after the first removal and approaches 0.5 when no secondary extinctions occur. When ranking species for removal, multiple species can sometimes receive the same rank value (a situation known as rank degeneracy). To address this, we employ a randomized tie-breaker. When ties occur, we randomly select which of the equally ranked species to remove. To account for the variability introduced by these random choices, we average the results from 200 repetitions of the entire removal process. In each repetition, ties are broken randomly. We use this approach to compare rankings based on our importance index, eigenvector in-centrality, and in-degree centrality. For eigenvector-based rankings, we follow the method in (Allesina and Pascual, 2009): at each step, we compute the dominant eigenvector of the column-normalized adjacency matrix and remove the highest-scoring node. The dominant eigenvalue is always 1, with its associated positive eigenvector summing to 1. For eigenvector-out rankings, we use the transpose of the adjacency matrix.

3.3 Results

3.3.1 Robustness index-Importance index Plane

Our algorithm allows us to place all the species within an ecosystem on the 2-dimensional space defined by the robustness index and the importance index. Each axis captures a different aspect of species, namely their importance as prey and carbon suppliers and their prowess as predators and carbon absorbers. As an example, we show in Fig. 6c the robustness index-importance index plane for the Cypress dry season food web (Ulanowicz, Bondavalli, and Egnotovich, 1998). In this figure, each point corresponds to a species, different colors/shapes correspond to five different categories of species, while point size is proportional to the biomass. Remarkably, species are not randomly scattered on the plane but tend to cluster following the classification. Producers and Organic Matter (dark green circles) are mostly placed on the top left corner since they are crucial carbon suppliers (high importance index) and they do not absorb carbon from other species (low robustness index). Conversely, large predators are placed in the bottom right corner, since they are both very hard to prey on and very good at preying. Also, a large fraction of secondary consumers are located in this region. The top right corner contains species that can absorb carbon from several other species, but that, at the same time, also transfer carbon to many other species. Not surprisingly in this area we find invertebrates and detritivores or mammals like rats. Finally, the bottom left corner mostly contains herbivores/omnivores and secondary consumers. These are species such as lizards that are not central in the food chain functioning, being hard to prey on, but with little to none preying capabilities.

In Fig. 7, the robustness index-importance index plane for each species is shown across all food webs. The plane is then divided into four regions:

- **Base Trophic Nodes (top left quadrant)**
  These species are characterized by poor predatory ability but are heavily preyed upon by a multitude of species. They form the foun-

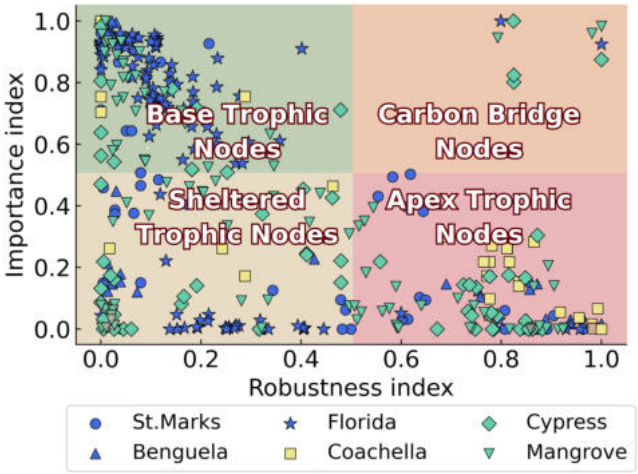

Figure 7: Robustness index-importance index plane. In royal blue, the marine ecosystems; in khaki, the terrestrial one; in medium aquamarine the mixed ones. Values are log-transformed and normalized to [0,1] for cross-food web comparison. The plane is divided into four quadrants. Base Trophic Nodes (top left quadrant), Carbon Bridge Nodes (top right quadrant), Sheltered Trophic Nodes (bottom left quadrant), Apex Trophic Nodes (bottom right quadrant).

dation for energy and nutrient flow, providing a primary source of food for consumers. They are both very important and very vulnerable (low robustness index).

- **Carbon Bridge Nodes (top right quadrant)**
  These species excel at absorbing or transforming carbon (often from detrital or non-living sources) and are widely consumed by other species. They "bridge" carbon from lower-level or non-living pools into higher trophic levels. They are also very important, but not as vulnerable, since their high robustness index allows them to absorb carbon from very diverse sources.

- **Sheltered Trophic Nodes (bottom left quadrant)**
  These species are neither strong predators nor frequent prey; they have defenses, behaviors, or ecological niches that protect them

from heavy predation. They have a low robustness index, so high vulnerability, but their extinction would likely not cause cascades.

- **Apex Trophic Nodes (bottom right quadrant)**
  These species possess high predatory ability with low vulnerability. They regulate prey populations and structure ecosystem dynamics from the top of the food web. They have a low importance index, but also low vulnerability.

Fig. 7 is realized considering six different ecosystems from terrestrial, aquatic, and mixed domains. Note that we normalized the robustness index and the importance index to compare species from different ecosystems. Colors denote the ecosystem type, while symbols show the different food webs. We observe a predominance of species in the Base Trophic Nodes (42.09%) and Apex Trophic Nodes (32.65%) quadrants, only a few Carbon Bridge Nodes (2.30%) and a minority of Sheltered Tropic Nodes (22.96%). We report in B.8, for each ecosystem, the detailed list of all species in each quadrant. It is important to stress that the robustness index and importance index only present a marginal anti-correlation (average Spearman $s = -0.60 \pm 0.18$ across the six food webs), confirming that the two quantities capture complementary species' properties. This moderate anti-correlation is notably lower than the correlation we observe between eigenvector in-centrality and importance index (e.g., $s = 0.93$ for the Cypress food web, as shown in B.3), indicating that the robustness index and importance index provide more distinct information.

3.3.2 Cascading Extinctions and Robustness of Species

The picture suggested by the robustness index-importance index plane, while reasonable, needs stronger validation. In particular, we need to understand whether the two measures we introduced are correlated with the predatory capabilities of animals and their vulnerability and importance as prey. These are not easy quantities to measure, but we can indirectly assess them.

We start studying the importance index, which, as mentioned, is a measure of a species' relevance as a carbon source. High-importance in-

Table 1: Comparison of mean extinction area across ecosystems using different species ranking methods. We compute in-degree, eigenvector-in, and importance index rankings at the outset and keep them fixed throughout the simulated extinction process. Values represent means $\pm$ standard deviation. When not shown, the standard deviation is smaller than the last significant digit displayed. Higher percentages indicate greater ecosystem sensitivity to species loss based on the respective ranking method. Bold values indicate the best-performing method for each ecosystem.

Food web	In-degree (%)	Eigenvector-in (%)	Importance index (%)
St. Marks	84.8 ± 1.1	**91.6**	91.4
Benguela	77.3 ± 0.5	**94.6**	91.3
Coachella	70.9 ± 0.3	75.9	**78.8**
Florida	64.6 ± 0.4	**82.4**	82.2
Cypress	78.5 ± 0.4	81.9	**86.9**
Mangrove	72.5 ± 0.1	**95.0**	94.3

dex species play a major role in food webs since they represent the base of the food chain. This role can be quantified looking at the extinction curve of a food web. The premise is that a more accurate measure of species relevance within a food web should lead to a faster network collapse when species are removed in order of their ranking. Indeed, Eq. (3.1) implies that a species with a high importance index is preyed upon by several species of low robustness index. Therefore, its removal would likely cause many of these species with low predatory capabilities to be unable to survive. We quantify this effect using the mean extinction area (explained in Sub Section 3.2.4), where higher values indicate a better ability to identify critical species within the ecosystem.

Table 1 presents the mean extinction areas following 3 different ranking strategies: in-degree, eigenvector-in centrality, and importance index. We establish the initial rankings and keep them fixed throughout the simulated extinction process (the "initial" attack variant described in Sub Section 2.3.3), whereas in B.4, we present the scenario where rankings are recalculated after each node removal (the "recomputed" attack variant), including extinction curves plots for both cases across all food webs. The results reveal that both the eigenvector-in and importance index consistently outperform the degree-centrality across all ecosystems

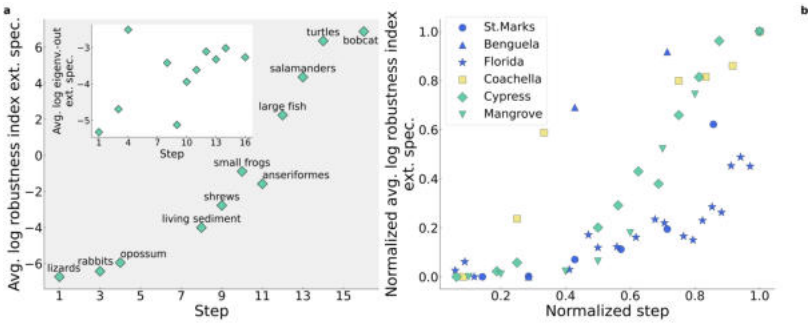

Figure 8: Ecological robustness: Species removal patterns in diverse food webs. Left: Cypress food web: Average log robustness index of extinct species vs. extinction step, based on importance index ranking. Labels indicate representative species becoming extinct at each step. **Inset:** Average log eigenvector-out centrality of extinct species vs. extinction step. **Right:** Normalized average log robustness index vs. normalized extinction step for all food webs. Values are log-transformed and scaled to [0,1] for cross-ecosystem comparison. This plot illustrates ecosystem resilience patterns and differential impacts of species loss across varied food web structures.

studied. Notably, eigenvector-in centrality and importance index demonstrate comparable efficacy, with each measure excelling in different ecosystems. These results validate the importance index, confirming its ability to identify the most crucial species within food webs.

As a second step, we study the ability of the robustness index to measure species' vulnerability. A high-robustness index species can prey on many other organisms, including those with a low importance index, i.e., those that are very hard to prey on. As a consequence, we expect high-robustness index species to be less affected by extinctions within the ecosystem. They are indeed very diversified and less susceptible to such events. On the other hand, low-robustness index species are expected to be the most vulnerable and, thus, the first to get extinct during cascade events. We can measure this by looking at the average (log) robustness index of species that get extinct at a given step of the extinction curve construction. As shown in Fig. 8a for the Cypress food web, this average increases with the step, meaning that high-robustness index species

only get extinct when most other species are already gone. We report next to each point a sample species going extinct in that step, while the inset shows the same plot, but using the average eigenvector-out centrality to measure (inverse) vulnerability. A much weaker correlation is observed. We repeat the same analysis for all the six food webs we considered, reporting the results in Fig. 8b. Note that we normalized both the step and the robustness index for comparison reasons. In all cases, we observe the same behavior. More quantitatively, we can compute the Spearman correlation between the removal step and the (average log) robustness index of the removed species. The average of this quantity among all ecosystems is 0.98, while the average of the eigenvector-out centrality is 0.72, thus consistently lower. We repeat the analysis using eigenvector-in centrality for species removal, yielding similar results that support the robustness index as an effective measure of species' predatory skills and extinction resilience. In B.5, we report details on Spearman correlation coefficients and eigenvector-in-based removals, along with additional comparisons between robustness index, importance index, and other centrality measures.

3.4 Discussion

Ecosystems face intensifying threats, making it urgent to identify vulnerable species, to prevent their extinction. Current network-based approaches often reduce the complex predator-prey interactions to a single metric. While this allows us to understand which species are crucial for the functioning of a food web, it tells little to nothing about the vulnerability of species. Our method, based on a bi-dimensional non-linear map, addresses this gap, associating to each species both an *importance index* and a *robustness index*. The former is linked to the susceptibility to predation and measures how crucial a species is in the food web. The latter measures predatory prowess and quantifies the robustness of species. High-importance index species, which often anchor lower trophic levels, appear critical to maintaining overall network stability, and their loss disproportionately impacts a wide range of other species. On the other

hand, high-robustness index species thrive on diverse prey pools, making them more resistant to ecosystem perturbations. Our analysis confirms this intuition, showing that the importance index effectively identifies species whose removal triggers significant co-extinctions. At the same time, the robustness index highlights species capable of persisting until the late stages of collapse, with the most vulnerable, low-robustness index species being the first to disappear. While eigenvector-in centrality can resemble our importance index ranking, neither eigenvector-in nor eigenvector-out centralities adequately capture the predatory strength and resilience that define a species' robustness index. This limitation arises from the linear nature of the eigenvector centrality, which is often incapable of capturing complex, non-linear interactions.

By mapping species onto a *robustness index-importance index plane*, we can reveal patterns that align with known functional roles (e.g., apex predators, base trophic nodes). Although expert knowledge remains crucial for fully characterizing species' functional roles, our algorithmic and data-driven approach provides a valuable, complementary layer of insight. This framework holds clear promise for conservation and ecosystem management, where limited resources demand strategic prioritization. By jointly quantifying species' importance index and robustness index, practitioners can pinpoint "keystone" nodes —those whose loss risks triggering wider co-extinctions— thereby directing scarce funding and protective measures where they have the greatest impact. Particular attention can also be devoted to those species that despite a marginal role as carbon suppliers, are in severe danger due to their low robustness index. Because our algorithmic approach draws solely on quantitative interaction data, it readily scales to large, complex food webs. When combined with objective risk assessment, this scalability ensures broader applicability across diverse ecosystems, ultimately supporting more targeted and cost-effective biodiversity preservation efforts. However, we acknowledge that the implementation of actual conservation strategies requires the consideration of numerous additional ecological, social, and economic factors beyond the scope of this network-based analysis.

While our framework provides valuable insights into species' roles

in food webs, certain aspects require further exploration. First, basal species are represented as individual nodes but are linked to a common root node that aggregates their shared reliance on environmental resources like light, water, and nutrients. This approach simplifies the analysis by not differentiating among their specific resource dependencies and instead focuses on their collective role as energy providers. Consequently, our analysis primarily examines the vulnerability and resilience of all other species in the food web. For basal species, however, additional factors such as Grime's classification (Grime, 2006)—based on competition, stress tolerance, and ruderality—may be necessary to better capture their ecological resilience and responses to cascading effects. Second, species such as lizards in the Cypress food web appear peripheral in our analysis due to their limited predatory skills and resistance to predation. This could reflect biological realities, such as population constraints driven by nesting site availability rather than trophic interactions. However, it may also result from aggregation biases in network construction, where diverse prey groups are aggregated into single nodes, masking their ecological importance. To preliminarily assess the influence of network resolution, we conducted node aggregation experiments. Our initial findings indicate that the impact of such aggregation on our metrics of ecosystem fragility is not uniform. Aggregating multiple basal trophic nodes yielded the most pronounced effects, suggesting that the resolution at lower trophic levels could be particularly critical for accurately assessing species' roles using our framework. Further details on these aggregation experiments can be found in B.6. Systematic investigation into the sensitivity of our results to food web resolution, across different aggregation strategies and trophic levels, represents a valuable direction for future research. Third, our analysis relies on a static and binary representation of food webs, which inherently simplifies the dynamic and often weighted nature of real ecological interactions. The species removal simulations, while useful for comparative purposes, also present an idealized view of extinction cascades, as they do not incorporate adaptive behaviors like prey switching or the continuous changes in interaction strengths that characterize natural ecosystems. Finally, future research

could readily expand this analysis to the larger collection of currently available food web datasets to assess the generality of our findings. Subsequently, experimental validation remains a critical avenue, where collecting data on real-world ecological networks under anthropogenic disturbances could provide empirical evidence for the predictive power of the proposed framework and help quantify species' actual endangerment levels across their ecosystems.

Chapter 4

Multilayer network analysis of European regional flows

4.1 Introduction

The economic vitality and developmental trajectory of European regions are increasingly defined not by their intrinsic attributes alone, but by their position within a complex web of inter-regional connections. This perspective aligns with the foundational concept of a *space of flows* (Castells, 1996), where the movement of capital, people, goods, and knowledge constitutes the fundamental architecture of the contemporary economy. In this relational view (Amin, 2004; Bathelt and Glückler, 2011), the prosperity and functional role of a region are contingent on its connectivity, i.e., the strength and diversity of its ties to other regions. While this phenomenon is often studied in the context of territorial attractiveness (Loris Servillo, Rob Atkinson, and A. P. Russo, 2012; A. Russo et al., 2013)—a crucial concept for integrating regional development strategies with the overarching goal of territorial cohesion (Faludi, 2006)—we argue that considering the underlying network dynamics is essential for a complete and nuanced understanding of regional development

and prosperity in today's integrated economy. Traditionally, research has attempted to map these interdependencies through two main lenses. The first analyzes individual flows in isolation, such as tourism (Cracolici and Nijkamp, 2009), migration (Waltert and Schläpfer, 2010), or investment (Jackson and Markowski, 1995). These analyses, while valuable, may provide an incomplete view that is unable to capture how different types of flows interact with one another. The second approach seeks a multidimensional view by creating composite indicators of territorial attractiveness (Musolino and Kotosz, 2024; OECD, 2022). However, these methods primarily treat regions as independent entities, aggregating their internal characteristics while largely overlooking the network structure of the flows that connect them and define their functional roles within the wider European system. We argue that a fundamental gap exists in understanding how different regional flows come together to form a cohesive system of multidimensional connectivity.

This chapter addresses this gap by proposing a paradigm shift from evaluating regional attractiveness as an inherent quality to analyzing multidimensional connectivity as a continuously interacting, relational process. We move beyond asking "How attractive is a region?" to asking "What is the structural role of a region within the interconnected European system?". To do this, we operationalize our framework using multilayer network science (Kivelä et al., 2014), modeling the European space as a multilayer network where each layer represents a distinct type of flow. This allows us to re-interpret what has been termed *revealed attractiveness* (Musolino, 2016)—not as a simple sum of inflows, but as a region's emergent structural position resulting from the complex interplay across multiple, co-existing networks. Different spatial scales can be used for this analysis, ranging from the macro-scale (e.g., countries (Reiner, Meyer, and Sardadvar, 2017)) to the micro-geographical scale (e.g., neighborhoods (Öner, 2017) or cities (Lee, 2016)). Regions at the NUTS 2 scale provide an ideal compromise, being neither too large to hide local dynamics nor too small to miss broader patterns of inter-regional connection. Our analysis of eight distinct flow types across European NUTS 2 regions from 2010 to 2018 is guided by the following

research questions:

R1 How does a region's importance transform when moving from an analysis of isolated flow types to a comprehensive multidimensional connectivity framework?

R2 What emergent functional relationships and regional clusters are revealed by the interplay of different flow types in a multilayer framework?

Our study first addresses the conceptual need to better capture the complex, interconnected nature of European regions by developing a multidimensional connectivity framework. This framework advances knowledge by moving beyond traditional single-dimension models, allowing for a richer understanding of how regions interact across flows of capital, knowledge, and people. Combining these datasets enables us to create a comprehensive, multidimensional view of regional connectivity, capturing how distinct flows of people, capital, goods, and knowledge collectively define the structure of the European space. This choice of layers is not arbitrary but is instead grounded in a substantial body of literature that identifies these specific flows as complementary channels of economic and social integration. Diaspora networks are known to reduce information costs and boost trade (Rauch and Trindade, 2002; Peri and Requena-Silvente, 2010; Docquier and Lodigiani, 2010); tourism has pro-trade effects (Santana-Gallego, Ledesma-Rodríguez, and Pérez-Rodríguez, 2011); passenger and air connectivity enable knowledge transfer and are associated with FDI (Ishutkina and Hansman, 2008; Fageda, 2017); and remittances are a primary measure of diasporic and financial links (World Bank and KNOMAD, 2023). Therefore, combining these distinct but interrelated flows in a multilayer network is an appropriate and well-established method for capturing a holistic view of connectivity (Kivelä et al., 2014; Boccaletti et al., 2014; Bonaccorsi et al., 2019). Operationally, we show that applying multilayer network techniques (i.e., multiplex PageRank centrality (Page, 1999) and Infomap community detection (Rosvall, Axelsson, and Bergstrom, 2009)) offers clear advantages over previous methods, as these tools can reveal intricate re-

gional structures and patterns that traditional economic or econometric methods may overlook (*IRiE Final Report* 2022; Komornicki, Rosik, and Mazur, 2023). Empirically, we demonstrate the value of this approach by providing a structurally aware understanding of European regional dynamics that identifies key hubs, tracks changes in regional roles over time, and uncovers functional ties that transcend national borders. Together, this conceptual, methodological, and empirical integration provides new insights into the evolving fabric of regional connectivity in Europe.

Our analysis reveals that European regional networks are characterized by a core–periphery structure, where a few regions dominate connectivity across multiple flow types. The multilayer perspective provides a more nuanced view of regional importance than single-layer analyses, highlighting the significant enhancement in the importance of regions like Bratislava and Leipzig. Furthermore, our community detection algorithm uncovers robust regional clusters, confirming the cohesive power of national borders but also revealing strong, unexpected cross-border functional regions. These findings underscore the necessity of a relational, multidimensional perspective for policymakers aiming to foster balanced regional development. The remainder of this chapter details the data and methodology used, presents the full results of our network analysis, and discusses the implications of our findings for regional science and European policy.

4.2 Materials and Methods

4.2.1 Data

ESPON is an EU-funded program providing territorial analyses, data, and maps. The dataset utilized in this study is derived from the IRiE ESPON project (*IRiE* 2022) and includes region-to-region (NUTS 2 level, 2016 version) origin–destination (OD) matrices covering various domains such as People Tourism, People Migration, Freight of Goods by transport mode, Capital Foreign Direct Investment (FDI), Knowledge (Erasmus

students), People Passengers by transport mode, Capital Remittances, and Knowledge (Horizon 2020). More details about the data, their processing, and the network construction can be found in C.1, while detailed information on the number of nodes, links, and density for each layer and year can be found in C.2.1.

4.2.2 Network Properties

We focus initially on single layers by analyzing first-order properties, such as the strength distribution, which helps us understand the basic structure of the networks by examining how strongly regions are connected. The out-strength (s_i^{out}) and in-strength (s_i^{in}) of node i are defined as follows:

$$s_i^{out} = \sum_j w_{ij} \quad \text{and} \quad s_i^{in} = \sum_j w_{ji}, \tag{4.1}$$

where w_{ij} represents the weight of the directed edge from node i to node j. We examine the complementary cumulative distribution function (CCDF) for in-strength, out-strength, and total strength.

Next, we move on to second-order properties, such as assortativity, which measures the tendency of nodes to connect to others that are similar or different in some way. Specifically, we focus on the Weighted Average Nearest Neighbors Strength (WANNS), i.e., $\text{WANNS}^{in,out}$, which is defined as follows:

$$\text{WANNS}_i^{in,out} = \frac{\sum_j w_{ij} s_j^{out}}{s_i^{in}}, \quad \forall i. \tag{4.2}$$

This calculates the weighted mean of the strengths of a node's neighbors.

4.2.3 Null Models

We use null models to validate the WANNS outcomes and determine whether the observed connections are due to underlying structural patterns or randomness. To achieve this, we compare our results with the CReMA null model (Parisi, Squartini, and Garlaschelli, 2020). This model

reconstructs the network topology and assigns weights to established links by maximizing entropy under given constraints. As long as these constraints are met (on average), all possible configurations are equally likely. This model is a specific instance of the continuous Directed Enhanced Configuration Model (Gabrielli et al., 2019). Specifically, enhanced configuration models (Squartini, Mastrandrea, and Garlaschelli, 2015) constrain sequences of both degrees and strengths. We use the NEMtropy package (Vallarano et al., 2021) to solve the model, employing the Newton method for both binary and weighted reconstructions, as well as the dcm-exp model for binary reconstruction, to generate an ensemble.

4.2.4 PageRank

To quantify regional importance, we first compute the PageRank centrality for single-layer networks (Page, 1999), then extend the analysis to a multi-dimensional perspective using multilayer PageRank via the muxViz package (version 3.1) in R (version 4.4.0) (De Domenico, Porter, and Arenas, 2015). In this study, PageRank serves as a proxy for regional importance by identifying regions that act as structural attractors across diverse socio-economic domains. Given that our dataset encompasses heterogeneous flows, simple strength-based metrics would only capture local popularity. In contrast, PageRank accounts for the recursive nature of systemic influence: a region is central not only if it receives significant flows, but if those flows originate from other systemically important regions. Unlike its single-layer counterpart, multilayer PageRank accounts for a random walker that can move both within layers and jump between layers. In this framework (De Domenico, Solé-Ribalta, et al., 2013), the system is modeled as a multilayer adjacency tensor, which represents a unified mathematical manifold where nodes and layers are treated as integrated components of a single state space. The dynamics are defined by a discrete-time random walk governed by a transition tensor derived from the row-normalized supra-adjacency matrix, ensuring that intra- and inter-layer transitions follow the same stochastic rules. Critically, this row-normalization prevents layers with higher total weights from

disproportionately biasing the global flow over layers with smaller units, as the exit probability from any node-layer pair is always conserved at unity. The relative probability of switching layers is thus locally balanced by the inter-layer coupling strength relative to the node's intra-layer degree. Consequently, PageRank serves as a unit-agnostic proxy for regional importance, identifying regions that act as systemic attractors across the European manifold based on the quality and transitivity of flows rather than raw volume.

4.2.5 Infomap Community Detection

Finally, we apply community detection to the multilayer network using the Infomap algorithm to reveal clusters of regions that exhibit strong functional coherence across multiple flow types. Unlike other methods that focus on topological density (e.g., modularity-based), Infomap is rooted in information theory and the Map Equation (Rosvall, Axelsson, and Bergstrom, 2009). It exploits the duality between finding community structure in a network and compressing the description of a random walker's trajectory through it. In this framework, a community is defined as a group of nodes in which a random walker remains for a relatively long time before escaping to another module. The algorithm optimizes the length of the codewords used to describe this trajectory by assigning shorter codewords to movements within modules and longer codewords to transitions between them, thereby identifying the optimal coarse-grained map of the system's flows. Key parameters in our implementation include the *two-level* setting for nested module detection, which identifies both broad regional blocs and more granular subclusters. Furthermore, we calibrate the multilayer relaxation rate, which governs the probability that the random walker switches layers. This parameter determines whether communities are determined primarily by individual flow types or by the integrated cross-layer structure. To ensure the stability of the identified clusters, we conducted a sensitivity analysis of the relaxation rate, assessing how varying the inter-layer movement probability affects the number of detected communities. De-

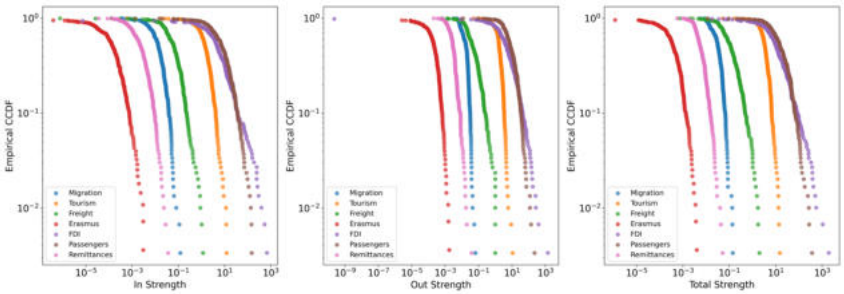

Figure 9: Complementary cumulative distribution function for the year 2010.

tailed information on the specific parameter selection, and the results of the sensitivity analysis are provided in C.4.1 and C.4.2.

4.3 Results

4.3.1 Single-Layer Networks

First-Order Properties

A key feature of many real-world networks is a heavy-tailed degree (or strength, in the weighted case) distribution. This general class of distributions signifies a consistent structural feature—the coexistence of a few nodes with exceptionally high connection strength and a large number of nodes with weak connectivity. To assess this pattern, we analyze the complementary cumulative distribution function (CCDF) of node strengths for the year 2010, as shown in Fig. 9. Indeed, all layers reveal a small number of regions acting as prominent hubs. Despite the somewhat limited range of these high-strength nodes, our results indicate that the networks exhibit this pattern, which persists in the 2018 data, as shown in C.2.2. This appendix subsection includes further analyses depicting the relationship between in-strength and out-strength for all flow types in 2010 and 2018.

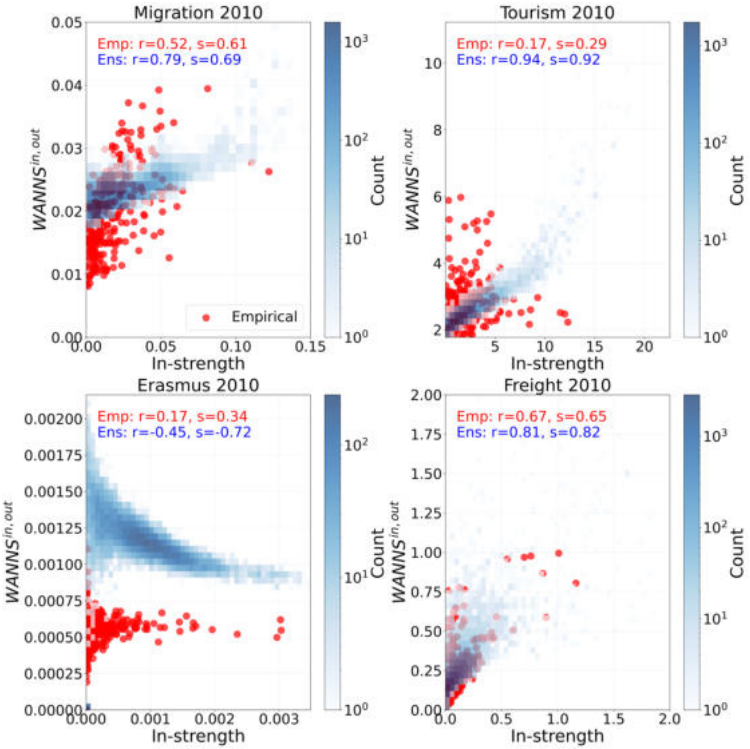

Figure 10: Empirical vs. ensemble WANNS for Migration, Tourism, Erasmus, and Freight (2010). Subplots show in-strength vs. WANNS^{in,out}. Red: empirical data; blue histogram: ensemble distribution. Pearson's r and Spearman's s correlations provided.

Second-Order Properties

Fig. 10 shows the Weighted Average Nearest Neighbor Strength (WANNS) for the empirical networks, alongside 50 realizations drawn from the null model ensembles for Migration, Tourism, Erasmus, and Freight. In C.2.3, we present the remaining three layers. The analysis reveals a predominantly assortative trend across all flow types. This assortative behavior indicates a positive correlation between node strengths and their neighbors' strengths, suggesting that regions with strong connectivity tend to

connect with other strongly connected regions. This analysis reveals the presence of a core–periphery structure, whereby the core is composed of regions that demonstrate a high level of interaction and exchange, while the periphery is constituted by regions with limited engagement in these flows. Comparing the empirical results with the null model predictions, we observe that the null model consistently anticipates a stronger correlation between WANNS^{in,out} and in-strength. This pattern holds true for all flow types except for Erasmus, where the null model predicts negative correlation coefficients, and for passenger transport, where both Pearson's and Spearman's correlations are lower in the null model, deviating from the trend observed in other flow types. Similarly, for FDI, we note that Spearman's correlation coefficient of the null model is smaller than the empirical one.

The finding that our empirical networks are less assortative than expected by random chance indicates that the European network is not as clustered as the distribution of its hubs would suggest. This finding challenges a simplistic view of a rigid core–periphery structure. It shows that hubs are also connecting to weaker, peripheral regions more often than predicted by the null model, suggesting a pattern of broader spatial integration and spillovers, despite the overall presence of a core–periphery structure.

Our approach provides a complementary perspective to methods, leveraging on exogenous inter-regional relationships that commonly use a pre-defined weighting matrix based on geographic distance or other contiguity measures (Anselin, 1988). While in such methods, spatial spillovers are generally assumed on a geographical basis, here, we define relationships endogenously, highlighting the functional geography of the system as dictated by the flows themselves. The assortativity that we observe is therefore not only a function of geographic proximity but it also suggests that functionally central regions connect to other functionally central regions, regardless of how far apart they are. This allows us to uncover the true interaction topology of the European system, identifying non-local, long-distance corridors of interaction that a standard spatial econometric model would miss.

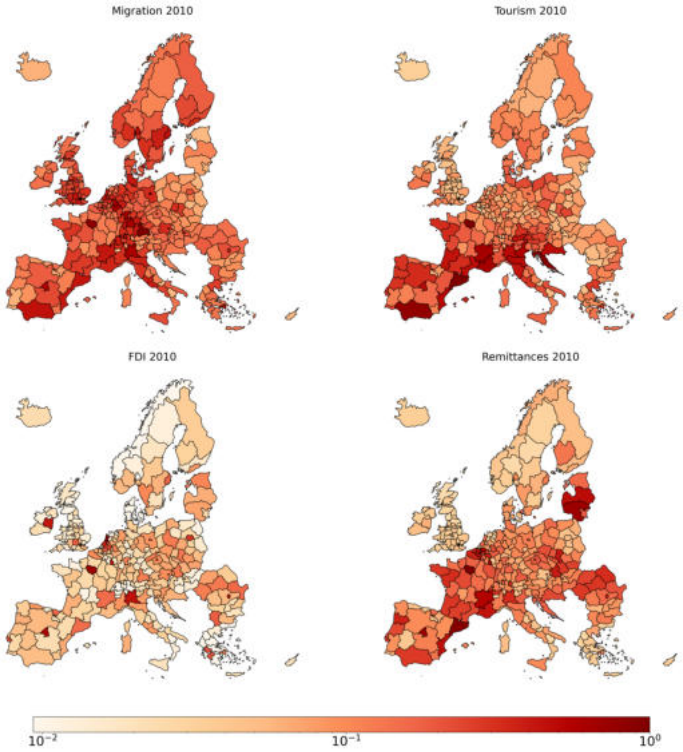

Figure 11: PageRank for Migration, Tourism, FDI, and Remittances in 2010. Colors are displayed on a logarithmic scale, with values normalized such that the region with the highest centrality is set to 1.

4.3.2 Centrality Measure: PageRank

Fig. 11 illustrates the spatial distribution of PageRank centrality values across European regions for Migration, Tourism, FDI, and Remittances in 2010. This visualization provides insights into the relative importance of regions within various flow networks, emphasizing the heterogeneity of regional centrality across different types of flows. In C.3, we show the remaining three layers, and in C.3.1, we present detailed tables showcasing the top 10 regions ranked by PageRank for all flow types in 2010.

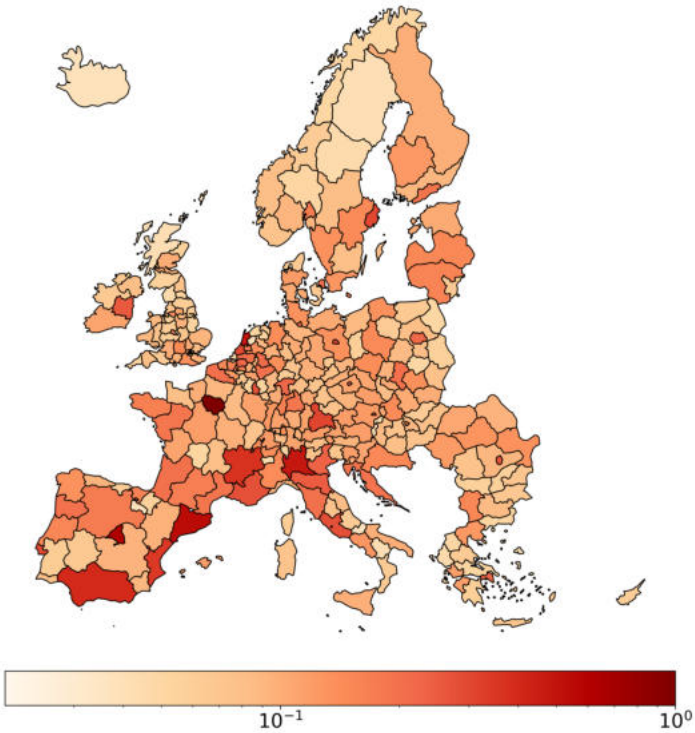

Figure 12: PageRank values for the 2010 multiplex. Colors are displayed on a logarithmic scale, with values normalized such that the region with the highest centrality is set to 1.

These rankings collectively demonstrate the varied roles that different regions play in different types of flows, with some regions, particularly Ile-de-France, showing high centrality across multiple networks. Furthermore, additional analyses on the relationship between PageRank and in-strength are included in C.3.2.

In contrast, Fig. 12 depicts the spatial distribution of multiplex PageRank centrality values for the same year, offering a comprehensive view of regional significance within the interconnected multilayer network structure by integrating information from all flow types.

Temporal Evolution

To capture the temporal evolution of PageRank rankings, in C.3.3 (Fig. 36–38) we show a heatmap of these rankings for each layer, with regions ordered according to their average position across all layers and years. The figures present a series of heatmaps organized in columns. Each column represents a specific type of flow, while the vertical axis provides the identifying name of the corresponding region. The width of each column is proportional to the number of temporal observations. The ranking's position is represented by shades of red, with higher saturation indicating a higher ranking position. This comprehensive visualization reveals a general trend of rank stability, particularly among top-ranked regions, which tend to maintain their positions across various flow types and years. However, we observe notable exceptions to this pattern, with certain regions demonstrating high centrality only in specific layers. Middle-ranked regions exhibit greater heterogeneity in their rankings across different flow types and years, indicating more dynamic centrality patterns in this tier.

Further analysis of PageRank trends (in the specific instance of the Migration layer) reveals that London maintains a strong upward trajectory in centrality through time, despite a temporary decline in 2016 (likely due to the Brexit referendum), highlighting the city's resilience as a key migration hub (see C.3.3 for details).

To distill key information from these temporal trends, Tables 2 and 3 present a focused analysis of PageRank ranking dynamics. We highlight regions with the highest and lowest average rankings, as well as those experiencing the most significant increases and decreases in ranking positions across layers. This analysis uncovers that certain regions, such as Ile-de-France, consistently maintain high centrality across multiple flow types, demonstrating their multifaceted importance in European networks. Conversely, other regions, like Lazio, exhibit exceptional centrality in specific domains, suggesting specialized roles within particular flow networks.

Table 2: Single-layer rankings (2010-2018), part I: highest and lowest average, largest increases and decreases for Erasmus, FDI, Freight, and Horizon2020 (excluding Melilla and Ceuta autonomous cities; Açores, Madeira, Åland, Canarias and Baleares island regions, and French overseas departments). E=East, N=North. Note: For FDI lowest average, other Danish regions, Inner London - East, and several Greek regions are omitted due to equal last position (zero inflows). Four Norwegian regions (Oslo og Akershus, Nord-Norge, Vestlandet, Trøndelag) are absent from the FDI network. For Erasmus largest increase and lowest average, regions absent from the network in some years are omitted from the rankings.

Category	Erasmus	FDI	Freight	Horizon2020
Highest Average: 1	Comunidad de Madrid (ES)	Noord-Holland (NL)	Lombardia (IT)	Ile-de-France (FR)
Highest Average: 2	Ile-de-France (FR)	Ile-de-France (FR)	Zuid-Holland (NL)	Comunidad de Madrid (ES)
Highest Average: 3	Andalucía (ES)	Luxembourg (LU)	Cataluña (ES)	Oberbayern (DE)
Highest Average: 4	Cataluña (ES)	Eastern and Midland (IE)	Ile-de-France (FR)	Région de Bruxelles (BE)
Highest Average: 5	Berlin (DE)	Comunidad de Madrid (ES)	Nord-Pas de Calais (FR)	Lazio (IT)
Lowest Average: 1	Sterea Ellada (EL)	Hovedstaden (DK)	Liechtenstein (LI)	Liechtenstein (LI)
Lowest Average: 2	Yugoiztochen (BG)	Thessalia (EL)	Ísland (IS)	Valle d'Aosta (IT)
Lowest Average: 3	Dytiki Makedonia (EL)	Espace Mittelland (CH)	Malta (MT)	Warmińsko-mazurskie (PL)
Lowest Average: 4	Valle d'Aosta (IT)	Ticino (CH)	Sostinės regionas (LT)	Cumbria (UK)
Lowest Average: 5	Warmińsko-mazurskie (PL)	Molise (IT)	Highlands and Islands (UK)	Opolskie (PL)
Largest Increase: 1	Kontinentalna Hrvatska (HR)	Bretagne (FR)	Kontinentalna Hrvatska (HR)	Extremadura (ES)
Largest Increase: 2	Jadranska Hrvatska (HR)	East Yorkshire (UK)	Ionia Nisia (EL)	Leipzig (DE)
Largest Increase: 3	Wielkopolskie (PL)	Zachodniopomorskie (PL)	Střední Morava (CZ)	Zentralschweiz (CH)
Largest Increase: 4	Podlaskie (PL)	Midi-Pyrénées (FR)	Jihovýchod (CZ)	Prov. West-Vlaanderen (BE)
Largest Increase: 5	Pomorskie (PL)	Yugoiztochen (BG)	Östra Mellansverige (SE)	Nord-Vest (RO)
Largest Decrease: 1	Nordjylland (DK)	Región de Murcia (ES)	N Ireland (UK)	Dél-Alföld (HU)
Largest Decrease: 2	Inner London E (UK)	Sardegna (IT)	Nyugat-Dunántúl (HU)	Basilicata (IT)
Largest Decrease: 3	Trentino (IT)	Dytiki Ellada (EL)	Peloponnisos (EL)	Moravskoslezsko (CZ)
Largest Decrease: 4	Sjælland (DK)	Lancashire (UK)	Dél-Alföld (HU)	Malta (MT)
Largest Decrease: 5	Liguria (IT)	Brandenburg (DE)	Etelä-Suomi (FI)	Jihovýchod (CZ)

Table 3: Single-layer rankings (2010-2018), part II: highest and lowest average, largest increases and decreases for Migration, Passengers, Remittances, and Tourism (excluding Melilla and Ceuta autonomous cities; Açores, Madeira, Åland, Canarias and Baleares island regions, and French overseas departments). E=East, W=West, NW=North West, NE=North East.

Category	Migration	Passengers	Remittances	Tourism
Highest Average: 1	Oberbayern (DE)	Ile-de-France (FR)	Ile-de-France (FR)	Jadranska Hrvatska (HR)
Highest Average: 2	Stuttgart (DE)	Inner London W (UK)	Cataluña (ES)	Cataluña (ES)
Highest Average: 3	Inner London E (UK)	Comunidad de Madrid (ES)	Luxembourg (LU)	Ile-de-France (FR)
Highest Average: 4	Darmstadt (DE)	Inner London E (UK)	Comunidad de Madrid (ES)	Andalucía (ES)
Highest Average: 5	Ile-de-France (FR)	Oberbayern (DE)	Rhône-Alpes (FR)	Rhône-Alpes (FR)
Lowest Average: 1	Liechtenstein (LI)	Liechtenstein (LI)	Liechtenstein (LI)	Liechtenstein (LI)
Lowest Average: 2	Valle d'Aosta (IT)	Ipeiros (EL)	Voreio Aigaio (EL)	Prov. Brabant Wallon (BE)
Lowest Average: 3	Molise (IT)	Burgenland (AT)	Ionia Nisia (EL)	Outer London W-NW (UK)
Lowest Average: 4	Malta (MT)	Corse (FR)	Highlands and Islands (UK)	Molise (IT)
Lowest Average: 5	Basilicata (IT)	Dytiki Makedonia (EL)	NE Scotland (UK)	Outer London E-NE (UK)
Largest Increase: 1	Vidurio ir vakaru Lietuvos (LT)	Etelä-Suomi (FI)	Ísland (IS)	Alentejo (PT)
Largest Increase: 2	Sud-Vest Oltenia (RO)	Brandenburg (DE)	Nyugat-Dunántúl (HU)	Zuid-Holland (NL)
Largest Increase: 3	Nord-Vest (RO)	Notio Aigaio (EL)	Pest (HU)	Jihovýchod (CZ)
Largest Increase: 4	Sud-Est (RO)	Leicestershire (UK)	Västsverige (SE)	Overijssel (NL)
Largest Increase: 5	Centru (RO)	Prov. Hainaut (BE)	Östra Mellansverige (SE)	Ísland (IS)
Largest Decrease: 1	Zentralschweiz (CH)	Yugozapaden (BG)	Attiki (EL)	Sud-Est (RO)
Largest Decrease: 2	Sicilia (IT)	Kontinentalna Hrvatska (HR)	Kentriki Makedonia (EL)	Lubuskie (PL)
Largest Decrease: 3	Campania (IT)	Yugoiztochen (BG)	Zuid-Holland (NL)	Dorset and Somerset (UK)
Largest Decrease: 4	Ostschweiz (CH)	Severoiztochen (BG)	Noord-Holland (NL)	Champagne-Ardenne (FR)
Largest Decrease: 5	Calabria (IT)	Latvija (LV)	Jadranska Hrvatska (HR)	Bourgogne (FR)

Table 4: Multiplex rankings (2010-2018): highest and lowest average, largest increases and decreases (excluding autonomous cities, Åland Islands, Atlantic island regions, and French overseas departments).

Category	Region
Highest Average: 1	Ile-de-France (FR)
Highest Average: 2	Comunidad de Madrid (ES)
Highest Average: 3	Noord-Holland (NL)
Highest Average: 4	Cataluña (ES)
Highest Average: 5	Lombardia (IT)
Lowest Average: 1	Liechtenstein (LI)
Lowest Average: 2	Molise (IT)
Lowest Average: 3	La Rioja (ES)
Lowest Average: 4	Flevoland (NL)
Lowest Average: 5	Voreio Aigaio (EL)
Largest Increase: 1	Bratislavský kraj (SK)
Largest Increase: 2	Leipzig (DE)
Largest Increase: 3	Alentejo (PT)
Largest Increase: 4	Kypros (CY)
Largest Increase: 5	Nord-Vest (RO)
Largest Decrease: 1	Dytiki Ellada (EL)
Largest Decrease: 2	Pohjois- ja Itä-Suomi (FI)
Largest Decrease: 3	West Central Scotland (UK)
Largest Decrease: 4	Northern Ireland (UK)
Largest Decrease: 5	Länsi-Suomi (FI)

Table 4 presents an analysis of the multiplex PageRank ranking across regions. It highlights regions with the highest and lowest average rankings, as well as those experiencing the most substantial positive and negative shifts in their ranking positions. The analysis reveals that Ile-de-France, Comunidad de Madrid, Noord-Holland, Cataluña, and Lombardia consistently maintain the highest average rankings in the multiplex network. This suggests that these regions play central roles across multiple types of flows within the European network. Conversely, we observe significant upward mobility in the rankings for regions such as Bratislava and Leipzig. These regions demonstrate the most substantial improvements in their multiplex PageRank positions, indicating an increase in their overall importance within the interconnected flow networks over time.

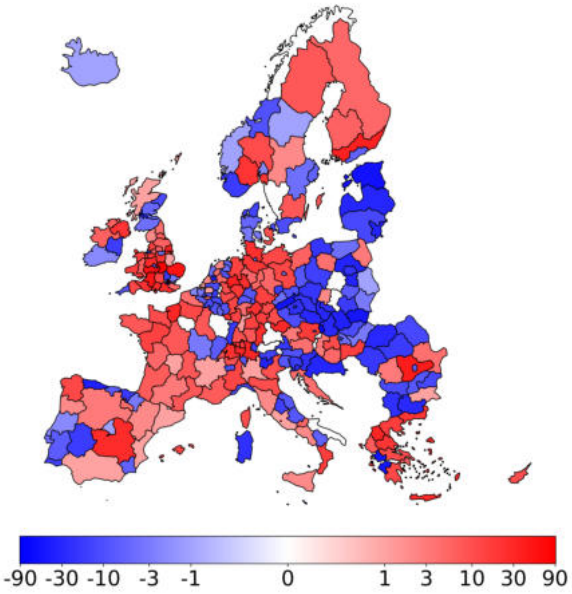

Figure 13: Change in node ranking: multilayer vs. average single-layer PageRank. Positive values (red) indicate an increase in ranking in the multilayer PageRank compared to the average of single-layer PageRanks, while negative values (blue) indicate a decrease. The intensity of the color represents the number of positions changed, with darker shades indicating larger changes. A logarithmic scale is used to emphasize changes near zero while still representing larger changes.

Single-Layer vs. Multiplex

To quantify the insights offered by the multilayer approach, we compare the regional rankings derived from the multiplex network against a simpler, aggregate baseline. This baseline is created by averaging the rankings of the single-layer PageRank for each region. Such single-layer averaging sets a baseline and, at the same time, represents a composite index of importance—a region's overall status is treated as the sum of its parts across different domains. The difference between a region's rank in the multilayer analysis and its rank in this simplified average baseline

is therefore a powerful indicator of multiplex effects. It reveals how a region's ability to integrate different flow types enhances or diminishes its overall importance in a way that simple aggregation cannot capture.

Fig. 13 illustrates the changes in node ranking when comparing these two approaches. A comprehensive explanation is provided below. The single-layer PageRank rankings were averaged to obtain a composite ranking, representing each node's average importance across all layers. The positions of nodes in the multilayer PageRank ranking were compared to their positions in the average single-layer PageRank ranking. The difference in ranking positions was calculated for each node. Nodes exhibiting an increase in ranking are indicated by positive values (red). This denotes an increase in the node's ranking in the multilayer PageRank relative to the average single-layer PageRank. That is, the node is more important in the multilayer analysis. Negative values (blue) indicate a reduction in the node's ranking within the multilayer PageRank relative to the average single-layer PageRank. This suggests that the node is of less importance in the multilayer analysis. This visualization reveals substantial variations, underscoring the importance of considering multiplex centrality measures to obtain comprehensive information not discernible from individual-layer analyses and revealing a heterogeneous pattern of ranking shifts. Notably, within individual countries, we observe both positive and negative shifts in regional rankings. The most consistent shift is observed for Malta, with a remarkable change of 91 positions, underscoring the potential for significant discrepancies between single-layer and multiplex centrality measures.

An analysis of the correlations between single-layer and multiplex PageRank (see C.3.4) reveals moderate positive relationships. While a perfect correlation would imply that the multilayer analysis is redundant, and no correlation would suggest that single-layer importance is irrelevant, the observed moderate correlation shows that while importance in a single, dominant flow contributes to a region's overall standing, it is not the sole determinant. The multilayer approach thus manages to capture the collective effect of diverse flow types, which is not fully visible from any single perspective.

Figure 14: Infomap community detection of European regional multiplex (2010). This map displays the NUTS 2 regions of Europe, with the different colors representing the 82 distinct communities detected by the Infomap algorithm applied to the multiplex network for the year 2010.

4.3.3 Community Detection: Infomap

The application of the Infomap algorithm to our multiplex network revealed a complex community structure across European regions (Fig. 14). This community detection algorithm identifies cohesive groups of regions based on the intensity of their mutual interactions across all flow types. A community represents a group of regions whose internal interactions are stronger than their connections to the rest of the network. These communities can be interpreted as empirically derived functional regions, i.e., economic and social subsystems whose boundaries are defined by flows, not by formal administrative lines. A total of 82 communities were identified, which appears to be a reasonable number given that the total number of regions and countries involved is approximately

300 and 30, respectively. This equates to an average of approximately two and a half communities per country. Our analysis reveals that these communities exhibit a mix of national cohesion and cross-border associations. This finding provides direct evidence for the dual nature of the European system. On the one hand, the enduring power of the nation-state remains a primary organizing force for many regional interactions. On the other hand, the emergence of strong cross-border communities signifies the development of transnational corridors of integration. The value of our approach lies in its ability to map the geography of this complex, multi-factor structure. Notable cases that emerged include the following:

- Belgium forms a community with Luxembourg and a neighboring Dutch region, suggesting strong economic and social ties in this cross-border area.

- The Czech Republic and Slovakia form a single community, reflecting their historical and ongoing close relations.

- A community comprises many English regions and Cyprus, indicating strong connections despite geographical distance.

- Several countries form predominantly self-contained communities, including Romania, Austria, Poland, Greece, Portugal, Hungary, Denmark, the Netherlands, Norway, Bulgaria, Finland, Malta, and Iceland. This suggests that these nations have stronger internal than external flows across the analyzed dimensions.

- Spain, France, and Italy each display a core community of multiple regions, with additional smaller communities, indicating complex internal structures.

- Cross-border communities are observed between Åland (Finland) and Sweden, as well as between Liechtenstein and Switzerland, highlighting strong regional ties that transcend national borders.

- The Baltic states (Lithuania, Estonia, and Latvia) form a cohesive community, reflecting their geographical proximity and shared historical background.

- Germany and the United Kingdom exhibit highly fragmented community structures, suggesting complex and diverse flow patterns within these countries.

- Slovenia, Croatia, and Malta form a community, potentially indicating strong economic or social ties among these Mediterranean and Adriatic regions.

- Northern Ireland and Ireland constitute a single community, aligning with their geographical proximity and historical connections.

Having 82 communities poses a set of challenges for their representation on a colored map. For more detailed information, we report the full listing of the communities in C.4.3.

4.4 Discussion

Our analysis, grounded in a multilayer network framework, contributes to several key discussions in regional science and economic geography. By moving beyond single-flow analyses, we provide empirical weight to the theoretical arguments for a relational understanding of regional economies (Amin, 2004; Bathelt and Glückler, 2011) and offer a more nuanced map of the *space of flows* within Europe (Castells, 1996).

A primary finding is the empirical validation of a core–periphery structure in Europe's functional network. The heavy-tailed distributions we observe are a common signature of many real-world complex systems (Albert and Barabási, 2002), and our analysis gives this property a concrete geographical meaning. The dominance of a few versatile hubs aligns with research on the world city network, which finds that a limited number of cities act as primary nodes for global capital and knowledge (Taylor and Derudder, 2015). Our work shows how this hierarchy

exists not just for global cities but also operates at the NUTS 2 regional level across a more diverse set of interactions.

The second key finding of our study is that moving from a single-layer to a multilayer centrality analysis fundamentally refines our understanding of regional importance. The primary contribution of the multiplex approach is its ability to look beyond a region's performance in one domain and assess its overall systemic role, distinguishing between two types of regions—specialized and versatile. While a single-layer analysis can clearly identify a region's specific strengths—for instance, the strong attractive power of German regions as specialized hubs for migration flows—it cannot, by itself, determine if this specialization translates into broader systemic influence. The multiplex perspective addresses this directly. We quantify this by measuring the rank shift, which is the change in a region's importance when moving from a simple average of its single-layer ranks to its integrated rank in the multilayer network. This shift is an indicator of non-linear, synergistic effects, revealing how a region's ability to combine diverse flows of capital, knowledge, and people enhances or diminishes its overall importance. This method yields significant empirical insights. The most pronounced ranking changes, for instance, are observed in Eastern Europe. This finding does not simply indicate growth, but rather points to a deep, structural transformation. The dramatic shifts suggest that the economic integration of these regions is not a shallow phenomenon, but a comprehensive evolution that is reshaping their role across the entire European system. This volatility means that as these regions integrate more deeply into the European system, they are actively sorting into distinct functional roles. Ultimately, the practical value of this approach lies in its ability to reveal the functional advantages of versatility over specialization, thereby providing a more realistic map of Europe's economic geography. It demonstrates that regional importance is a multi-faceted characteristic, not a single, unified concept. The ability to effectively bridge different economic and social dimensions constitutes a distinct and crucial form of centrality that conventional analyses, which examine each flow independently, overlook. This offers a new lens to policymakers,

presenting a strategic choice between development strategies that reinforce a successful specialization and those that foster the network versatility that appears to be a key feature of systemically important regions. These general findings are further substantiated by specific case studies. Notably, Bratislava and Leipzig demonstrate the most significant increases in multiplex ranking from 2010 to 2018. This finding aligns with their documented economic trajectories during this period. Bratislava, despite a relatively stagnant population, experienced remarkable economic growth (Hanzl-Weiss, Holzner, and Römisch, 2018). Its GDP per capita at purchasing power parity surpassed that of Vienna, placing it among Europe's top 10 leading regions. This economic surge was primarily driven by substantial foreign direct investment, particularly in the automotive sector, leading to full employment in the region. Similarly, Leipzig emerged as Germany's fastest-growing city in the 2010s (Cudny and Kunc, 2022). Its remarkable growth can be attributed to massive public investments, subsidies, and support programs across various policy fields and sectors. These public initiatives were instrumental in mobilizing significant private capital investments across all urban sectors, fueling the city's rapid development. These case studies of Bratislava and Leipzig illustrate how the multiplex analysis captures complex regional dynamics that might be missed in single-layer examinations. The multiplex approach effectively reflects the multifaceted nature of regional development, encompassing factors such as foreign investment, economic growth, and policy interventions, which collectively influence a region's centrality within the European network of flows.

Finally, our community detection results provide direct, empirical evidence on the long-standing debate between European integration and national cohesion, revealing a complex reality that is not an either/or scenario. On the one hand, the emergence of strong cross-border communities validates the concept of a Europe of regions, where functional economic units are defined by flows rather than formal borders. The integrated community spanning the island of Ireland or the cohesive block linking Belgium, Luxembourg, and a Dutch region are prime examples of this phenomenon, aligning with scholarly work on the rise

in global city-regions (Scott, 2001). On the other hand, our analysis simultaneously confirms the enduring power of the nation-state. The persistence of strong, self-contained national communities, even in highly globalized countries like Denmark and the Netherlands, demonstrates that the state remains the primary container for a dense web of socio-economic flows. Together, these findings depict a European space characterized by a dual structure—a mosaic of deeply integrated national systems overlaid with powerful, transnational corridors of interaction. Beyond simply transcending administrative borders, our flow-based approach challenges an even more fundamental assumption—the primacy of geographic distance. By revealing non-intuitive communities, such as the link between English regions and Cyprus, the analysis demonstrates that intense relational proximity (e.g., through strong financial ties) can be far more significant for community formation than spatial proximity. More specifically, the England–Cyprus community result is no longer counter-intuitive when viewed through a non-geographic lens. The connection is a direct reflection of relational ties, most notably the strong, historic financial links between the UK and Cyprus, which is a major financial hub, and their shared status as Commonwealth members. The algorithm correctly identifies that the combined strength of these specific connections creates a more cohesive community than the ties between those English regions and many of their geographically closer European neighbors. This finding offers a contrast to methodologies like spatial econometrics, which often presuppose that distance is the key determinant of interaction. By letting the data define the connections, our work uncovers the functional topology of Europe and provides an empirical method for analyzing the relational assets of regions, which is a central task for contemporary economic geography. This growing understanding of relational proximity over distance not only reveals unexpected cross-regional connections but also sheds light on situations where historical and political divisions might suggest limited integration. Our results highlight dynamics that could remain hidden when using only traditional perspectives. For example, on the island of Ireland—an area marked in recent history by political conflict and border tensions—our

flow-based analysis shows that Ireland and Northern Ireland function as a single integrated community. Although a purely historical or political lens might emphasize division, the empirical patterns in cross-border flows reveal a deep economic and social interdependence on the island. This means that even where legacy narratives predict separation, everyday connections can point towards unexpected forms of integration. Thus, our approach refines our understanding of territorial cohesion and serves as a tool to uncover the latent patterns of interaction throughout Europe.

Chapter 5

Conclusion

This thesis embarked on a journey to expand the analytical power of network science, driven by the principle that our understanding of complex systems is only as good as the tools we use to study them. We began by identifying key limitations in established methodologies and then proceeded to develop, apply, and extend network-based frameworks to gain deeper insights into ecological and socio-economic systems. The research was motivated by two primary challenges. First, we confronted the architectural constraints of influential bipartite ranking algorithms, such as the Economic Complexity Index (ECI) and the Economic Fitness and Complexity (EFC) algorithm, whose applicability has been historically confined to a narrow class of networks. Second, we addressed the inherent inadequacy of single-layer analyses for capturing the multifaceted nature of complex socio-economic systems, where different types of interactions overlap and influence one another. This concluding chapter synthesizes the primary contributions of the thesis in addressing these challenges. It will discuss the broader implications of our findings, acknowledge the limitations of the research, and finally, propose promising directions for future inquiry.

A central contribution of this thesis was a significant methodological advancement: the generalization of the Economic Complexity Index (ECI) and the Economic Fitness and Complexity (EFC) algorithm

for application to monopartite networks. This breakthrough bridged a crucial gap, extending the applicability of these non-linear methods beyond their limitation to strictly bipartite structures. The primary outcome was the development of fitness centrality, a novel measure proven to be highly effective at identifying "crucial" nodes—those connected to many nodes with a low degree—offering a complementary perspective to traditional centrality metrics. This methodological innovation immediately enabled a new approach to a pressing ecological problem. Recognizing that single metrics often fail to capture the dual roles species play, we applied our framework to develop a bi-dimensional model for food webs. This method characterizes each species by an importance index (its role as a carbon source and susceptibility to predation) and a robustness index (its predatory prowess and resilience). Our findings confirmed that this decoupled approach provides a far more nuanced view of ecosystem stability. The importance index effectively identifies species whose removal triggers significant co-extinctions, while the robustness index successfully predicts which species are most vulnerable to collapse–a feature not captured by standard linear measures like eigenvector centrality. Expanding from the analysis of single networks, the final part of this thesis addressed the complexity inherent in multifaceted socio-economic systems. We constructed and analyzed a multilayer network of European regions, demonstrating that this integrated perspective offers a more structurally aware understanding than single-layer analyses. The key contribution was the ability to distinguish between specialized regions, dominant in a single domain, and versatile regions, whose importance stems from their ability to bridge multiple types of flows. The dramatic rise in importance of regions like Bratislava and Leipzig in the multilayer view exemplifies this finding, revealing a new dimension of regional strength rooted in systemic integration rather than sectoral dominance.

The contributions of this thesis extend beyond their immediate contexts, offering broader implications for network science, conservation ecology, and regional policy. For the field of Network Science, this work provides a significant expansion of the analytical toolkit. By generaliz-

ing the ECI and EFC algorithms, we have transformed them from specialized methods for bipartite systems into versatile tools applicable to different network structures. The most significant outcome of this is the introduction of fitness centrality as a novel measure of node importance. Unlike degree or eigenvector centrality, which measure popularity or connection to popular nodes, fitness centrality offers a complementary perspective by efficiently identifying the nodes whose removal leads to the isolation of many others. This provides a computationally inexpensive yet powerful method for assessing a specific and critical type of network vulnerability, offering a new lens through which to understand structural integrity. For Ecology and Conservation Science, this methodological advance translates into a tangible and scalable tool for ecosystem management. By mapping species onto a robustness-importance plane, our framework moves beyond single-metric assessments. This allows conservation practitioners to strategically prioritize efforts with limited resources. They can identify not only systemically important "keystone" species whose loss would trigger co-extinctions but also those species that, despite a marginal role, are highly vulnerable due to low robustness. As an algorithmic approach based on quantitative data, this method is readily scalable to large food webs, supporting more targeted and cost-effective biodiversity preservation efforts worldwide. Finally, the principles of integrated network analysis carry significant implications for Regional Science and Public Policy. Our multilayer analysis of the European space provides a more realistic and structurally aware model of regional dynamics. The key distinction between versatile and specialized regions offers a powerful tool for tailoring development strategies—fostering diversification in the latter and leveraging the bridging role of the former. Furthermore, the identification of robust, cross-border functional communities suggests that cohesion policies could be more effective if targeted at these empirically defined regions rather than being constrained by administrative boundaries. This provides policymakers with an evidence-based lens to better understand and navigate the interconnected reality of the European economy, potentially leading to a greater return on investment for integration initiatives.

While the findings of this thesis offer significant contributions, it is essential to acknowledge their limitations, which provide a foundation for future inquiry. These limitations span the core methodology, the specific applications, and the data upon which they rely. Methodologically, while the generalization of ECI and EFC provides a valuable new tool, its theoretical properties are not yet fully understood. We observe a consistent power-law-like relationship between fitness centrality and degree across various networks, but we currently lack a clear theoretical explanation for this phenomenon. Furthermore, the iterative algorithm exhibits non-uniform convergence properties; while high-fitness nodes can be estimated rapidly, achieving convergence for low-fitness nodes is computationally more demanding. In its application to ecological systems, the framework relies on a static and binary representation of food webs, which inherently simplifies the dynamic and weighted nature of real-world interactions. The model does not incorporate adaptive behaviors like prey switching, and it simplifies the role of basal species by aggregating their diverse resource dependencies. Our preliminary experiments confirm that the framework is sensitive to the resolution of the network data, particularly how species at lower trophic levels are aggregated, which warrants a more systematic investigation. Similarly, the analysis of European regions is constrained by the nature of the available data. The results are contingent on the quality and consistency of the ESPON dataset, and potential biases in data collection or harmonization could influence the observed network structures. Critically, while the multilayer network reveals structural correlations, it does not establish causation. The findings are therefore best understood as a data-driven map that generates new, specific hypotheses for future qualitative and econometric investigation, which would be needed to explain the causal mechanisms driving the observed connections.

The findings and limitations of this thesis naturally point toward several promising avenues for future research, spanning methodological refinement, empirical validation, and the integration of new socio-economic dimensions. First, future work could focus on refining the core methodology introduced in Chapter 2. The generalized fitness map could be

modified by adjusting its non-linear term, for example, by introducing a parameter ν to amplify or diminish the importance of low-degree nodes. As suggested by related work in the bipartite case, this could allow for fine-tuning the algorithm's correlation with degree centrality. Second, a critical next step is to test the robustness and generality of our findings in real-world contexts. The ecological framework from Chapter 3 should be applied to the larger collection of available food web datasets to assess the universality of the observed patterns. More importantly, experimental validation is needed. Collecting data from ecosystems undergoing anthropogenic disturbances could provide direct empirical evidence for our framework's predictive power, helping to quantify species' actual endangerment levels. Finally, the multilayer framework from Chapter 4 offers a foundation for exploring the drivers and consequences of the observed network structures. For example, the high multilayer centrality of regions like Île-de-France, Madrid, and Lombardy corresponds strongly with their role as hubs for the economic integration of stable, non-EU immigrant populations[1][2][3]. This suggests a compelling hypothesis for future investigation: that an integrated immigrant workforce provides a "system elasticity," easing labor shortages and creating a foundation upon which high-productivity sectors thrive, in turn boosting the very flows that define a region's systemic importance. Testing this and other relationships between network structure and socio-economic phenomena represents a rich direction for future research, moving from describing complex systems to explaining them.

In summary, this thesis has progressed from methodological innovation to real-world application. It began by generalizing a class of powerful but restricted algorithms, creating new analytical tools for the network science community. It then applied these tools to address pressing challenges in ecology, demonstrating their practical utility in under-

[1] https://www.immigration.interieur.gouv.fr/Info-ressources/Etudes-et-statistiques/Les-chiffres-de-l-immigration-en-France/Population-immigree-par-departement? accessed on 16/09/2025.

[2] https://www.ine.es/dyngs/Prensa/es/ECP1T24.htm accessed on 16/09/2025.

[3] https://www.istat.it/wp-content/uploads/2024/10/REPORT-CITTADINI-NON-COMUNITARI_Anno-2023.pdf? accessed on 16/09/2025.

standing ecosystem stability. Finally, it employed advanced multilayer techniques to reveal the hidden, multifaceted structure of the European economy. Together, these contributions underscore the power of a versatile and integrated network science approach to decode the intricate systems that shape our world and provide a more robust foundation for evidence-based policy.

Appendix A

Supplementary Information Chapter 2

A.1 Data

In the following subsections, we describe the networks we use to benchmark our method. All of the networks can be downloaded from Kunegis (2013).

A.1.1 The Zachary Karate Club

The Zachary Karate Club network (ZKC) was sourced from data gathered by Wayne Zachary in 1977 from the members of a university karate club (*Zachary karate club network dataset – KONECT* 2017; Zachary, 1977). In this network, each node stands for a club member, and each edge signifies a connection between two members. The network is undirected with 34 nodes and 78 links. This dataset is often used as a test-bed for community detection algorithms. It involves identifying the two factions that emerged within the karate club following a dispute between two instructors.

A.1.2 Dolphins

This undirected social network represents a community of bottlenose dolphins (genus *Tursiops*) living in Doubtful Sound, a fjord in New Zealand (*Dolphins network dataset – KONECT* 2017; Lusseau et al., 2003). Each node corresponds to a dolphin, and an edge indicates a frequent association between two dolphins. The observations were conducted between 1994 and 2001, resulting in a network of 62 nodes and 159 links.

A.1.3 Train bombing

This undirected network depicts the contacts between suspected terrorists involved in the Madrid train bombing on March 11th, 2004, as reconstructed from newspapers. Each node represents a terrorist, and an edge between two nodes indicates contact between them. The edge weights represent the strength of the connection, including friendships and coparticipation in training camps or previous attacks, although we considered the network as unweighted. The network comprises 64 nodes and 243 links.

A.1.4 Jazz musicians

This is the collaboration network between Jazz musicians (*Jazz musicians network dataset – KONECT* 2017; Gleiser and Danon, 2003). Each node is a Jazz musician and an edge denotes that two musicians have played together in a band. The data was collected in 2003. This network contains 198 nodes and 2,742 links.

A.1.5 Caenorhabditis elegans

This is a weighted, directed network representing the neural network of *Caenorhabditis elegans* (*Caenorhabditis elegans (neural) network dataset – KONECT* 2018; White et al., 1986). The original network featured directed edges, allowing multiple parallel edges with integer weights. In the version we use in the main text, the network is undirected and unweighted. We use the undirected weighted version of this network in

Fig. 21 to demonstrate that fitness can be easily calculated in weighted networks. As described in the original publication by J. G. White et al., the data was assembled manually and may contain a small number of errors. To cite:

> We are reasonably confident that the structure that we present is substantially correct and gives a reasonable picture of the organization of the nervous system in a typical *C. elegans* hermaphrodite.

This network consists of 297 nodes and 4,296 links.

A.1.6 US airports

This network depicts flight connections between US airports in 2010, using a directed weighted graph (*US airports network dataset – KONECT* 2017; Opsahl, 2011). Each edge indicates a specific airport route, with the edge's weight representing the total number of direct flights between those airports in the given direction. In Fig. 21, we use the undirected weighted version of this network to test the fitness algorithm in weighted networks. This network consists of 1,574 nodes and 28,236 links.

A.2 ECI spectral analysis

Existing literature indicates that ECI is connected to the spectral properties of specific matrices (Caldarelli et al., 2012; Mealy, Farmer, and Teytelboym, 2019). In this section, we will provide a brief overview of the spectral method used for estimating ECI. We begin with the original ECI map from Eq. (2.2) and write it as

$$\begin{cases} \vec{F}^{(n)} = \mathbf{D}^{-1}\mathbf{M}\,\vec{Q}^{(n-1)} \\ \vec{Q}^{(n)} = \mathbf{U}^{-1}\mathbf{M}^{T}\,\vec{F}^{(n-1)}, \end{cases} \tag{A.1}$$

with the matrices $\mathbf{D}^{-1}$ and $\mathbf{U}^{-1}$ defined as square diagonal matrices, where the diagonal elements are the inverses of country diversification and product ubiquity, respectively. These matrices have dimensions $N_c \times$

N_c and $N_p \times N_p$, respectively. The quantities $\vec{F}^{(n)}$ and $\vec{Q}^{(n)}$ are respectively N_c and N_p dimensional vectors and the superscript T represents matrix transposition. After a single iteration, a step that separates the evolution of economic complexity from that of product complexity, we get:

$$\begin{cases} \vec{F}^{(n)} = \mathbf{D}^{-1}\mathbf{M}\mathbf{U}^{-1}\mathbf{M}^T \ \vec{F}^{(n-2)} \\ \vec{Q}^{(n)} = \mathbf{U}^{-1}\mathbf{M}^T\mathbf{D}^{-1}\mathbf{M} \ \vec{Q}^{(n-2)}, \end{cases} \tag{A.2}$$

We then define the matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ as $\mathbf{S}_1 = \mathbf{M}\mathbf{U}^{-1}\mathbf{M}^T$ and $\mathbf{S}_2 = \mathbf{M}^T\mathbf{D}^{-1}\mathbf{M}$. We notice that these matrices are symmetric, indeed

$$\begin{cases} \mathbf{S}_1^T &= \left(\mathbf{M}\mathbf{U}^{-1}\mathbf{M}^T\right)^T = \mathbf{M}\mathbf{U}^{-1}\mathbf{M}^T = \mathbf{S}_1 \\ \mathbf{S}_2^T &= \left(\mathbf{M}^T\mathbf{D}^{-1}\mathbf{M}\right)^T = \mathbf{M}^T\mathbf{D}^{-1}\mathbf{M} = \mathbf{S}_2, \end{cases} \tag{A.3}$$

since $\mathbf{U}^{-1}$ and $\mathbf{D}^{-1}$ are diagonal. Finally, defining the matrix $\mathbf{N}_1$ as $\mathbf{N}_1 = \mathbf{D}^{-1}\mathbf{S}_1$ and the matrix $\mathbf{N}_2$ as $\mathbf{N}_2 = \mathbf{U}^{-1}\mathbf{S}_2$, we can express Eq. (A.2) as

$$\begin{cases} \vec{F}^{(n)} = \mathbf{N}_1 \ \vec{F}^{(n-2)} \\ \vec{Q}^{(n)} = \mathbf{N}_2 \ \vec{Q}^{(n-2)}. \end{cases} \tag{A.4}$$

In terms of the initial conditions, we have

$$\begin{cases} \vec{F}^{(2n)} = \mathbf{N}_1^n \ \vec{F}^{(0)} \\ \vec{Q}^{(2n)} = \mathbf{N}_2^n \ \vec{Q}^{(0)}. \end{cases} \tag{A.5}$$

It is worth noting that if both matrices $\mathbf{U}$ and $\mathbf{D}$ were identity matrices, meaning economic and product complexities were defined as sums instead of averages, then $\mathbf{N}_1$ and $\mathbf{N}_2$ would become Kleinberg matrices. These matrices were initially employed to calculate authority and hub scores during the early days of the World Wide Web (Kleinberg, 1999), providing an alternative to the PageRank algorithm (Brin and Page, 1998). In our context, both $\mathbf{N}_1$ and $\mathbf{N}_2$ are transition matrices, also called normal matrices. This stems from the multiplication of $\mathbf{M}$ and $\mathbf{M}^T$ by $\mathbf{D}^{-1}$ and $\mathbf{U}^{-1}$, respectively, which normalizes their rows, ensuring that the elements in each row add up to one. Normal matrices possess real eigenvalues within the range of -1 to 1. The principal eigenvector, associated with the eigenvalue one, is a constant non-null vector. Consequently,

the map defined in Eq. (A.5) is destined to converge to a vector with constant components, namely the principal vector. Therefore, when calculating ECI, the process must be halted after a certain number of iterations to obtain meaningful results. This approach is akin to considering a linear combination of the first non-trivial eigenvectors, with the one corresponding to the second largest eigenvalue being the most relevant (Caldarelli et al., 2012).

Considering the matrix $\mathbf{N}_1$ (similar reasoning can be applied to $\mathbf{N}_2$), its eigenvectors $\vec{\varphi}_i$ with corresponding eigenvalues λ_i form a complete (non-orthogonal) basis in the space of N_c dimensional vectors. Organizing the eigenvalues λ_i in descending order, where $\lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_{N_c}|$, we can express $\vec{F}^{(0)}$ as:

$$\vec{F}^{(0)} = a_1\vec{\varphi}_1 + a_2\vec{\varphi}_2 + \cdots + a_{N_c}\vec{\varphi}_{N_c}. \tag{A.6}$$

Further, rewriting the first equation of Eq. (A.5) gives:

$$\begin{aligned} \vec{F}^{(2n)} &= a_1\vec{\varphi}_1 + a_2\lambda_2^n\vec{\varphi}_2 + \cdots + a_{N_c}\lambda_{N_c}^n\vec{\varphi}_{N_c} \\ &= a_1\vec{\varphi}_1 + a_2\lambda_2^n\vec{\varphi}_2 + \lambda_2^n O((\lambda_3/\lambda_2)^n). \end{aligned} \tag{A.7}$$

Consequently, for sufficiently large n, the country ranking is determined by the components of $\vec{\varphi}_2$, since $\vec{\varphi}_1$ is constant. Hence, computing the eigenvector components associated with the second eigenvalue reveals the country ranking.

A.3 Correlations between fitness and other centrality measures

A.3.1 Real-world networks: betweenness, eigenvector, and degree centralities

As stated in the main paper, a power-law relationship can be observed between the centrality measures and fitness. To quantify this relationship, the exponent of fitness is estimated using least squares regressions, as illustrated in Table 5. Additionally, Spearman rank correlation coefficients are included in the table's rightmost column.

Table 5: Power-law dependence on fitness centrality. Assuming a power-law relationship between fitness and the betweenness, eigenvector, and degree centralities, we estimated the exponents using a least-squares fitting procedure. The resulting exponent values are reported in column 3 and their R^2 values are in column 4. The Spearman rank correlation coefficients between these quantities are also provided in the last column.

Centrality Measure	Network	Exponent	R^2	Spearman Corr.
Betweenness	Zachary	0.37	0.33	0.80
	Train	0.85	0.60	0.80
	Jazz	1.95	0.61	0.86
	C. elegans	0.86	0.47	0.81
	Dolphins	0.51	0.24	0.79
Eigenvector	Zachary	0.1	0.30	0.47
	Train	0.37	0.26	0.55
	Jazz	1.03	0.47	0.74
	C. elegans	0.3	0.37	0.54
	Dolphins	0.3	0.19	0.42
Degree	Zachary	0.17	0.78	0.88
	Train	0.38	0.80	0.89
	Jazz	0.79	0.81	0.90
	C. elegans	0.42	0.81	0.87
	Dolphins	0.31	0.75	0.87

A.3.2 Real-world networks: ECI

Figure 15 reveals no significant correlation between ECI and fitness centrality. This lack of correlation is expected, as these metrics are fundamentally designed to capture distinctly different network characteristics. The Spearman rank correlation is less than 0.22 in all the five networks studied.

A.3.3 Synthetic networks

Figure 16 illustrates the correlation between fitness and other centrality measures (degree, eigenvector, and betweenness) for distinct synthetic network types (Erdős–Rényi, Barabási-Albert, and Watts-Strogatz). In the case of a random graph, all measures are expected to correlate with degree and thus with each other, as there is no inherent structure to the network. In the case of Barabási-Albert and Watts-Strogatz, there is a rel-

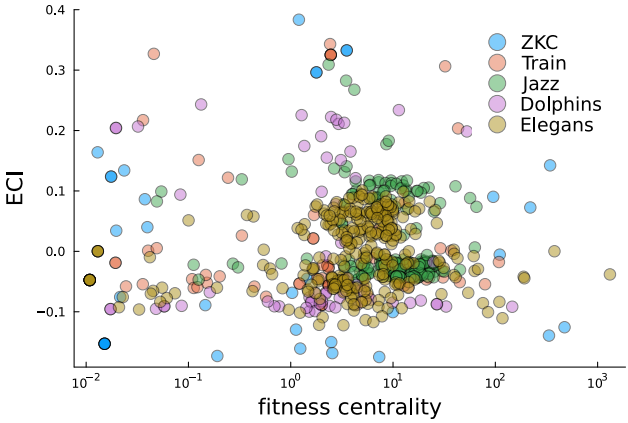

Figure 15: ECI vs. fitness centrality. The plot demonstrates that the correlation between ECI and fitness centrality is negligible across all five networks analyzed. The Spearman rank correlation remains below 0.22 in every case.

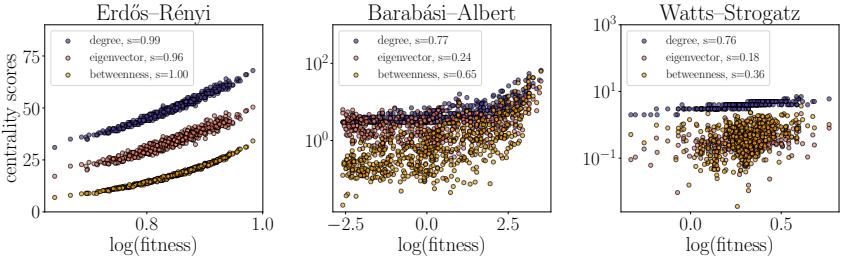

Figure 16: Centrality measures correlations in synthetic networks. We observe perfect correlations among all centrality measures in random Erdős-Rényi networks (we report the s Spearman rank correlation s in the legends). For Barabási-Albert and Watts-Strogatz networks, high correlation is observed only with the degree, while correlations with other centrality measures range from low to moderate. All networks are generated with $N = 500$ nodes. The parameters used are as follows: for ER, $p = 0.1$; for BA, $m = 3$; for WS, $k = 4$, $p = 0.1$.

atively weak correlation between the fitness and both betweenness and eigenvector centrality. This suggests that the variable is extracting a different set of information.

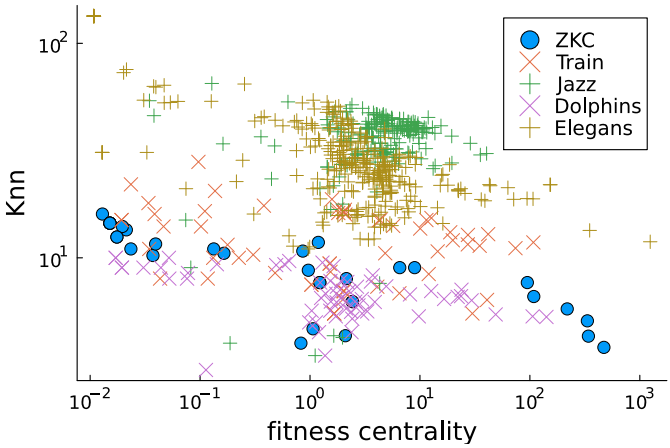

Figure 17: Dependence between average nearest neighbor connectivity (Knn) and fitness centrality for real-world networks. We report the least square estimate of the possible power-law dependence and the corresponding R^2 values in Table 6. We find a mild anti-correlation in the case of ZKC and Elegans, with $R^2 = 0.66$ and 0.45 respectively, and a low R^2 values in the remaining three cases. The Spearman rank correlation in the last column agrees with the previous consideration except in the case of ZKC where the Spearman correlation is more prominent.

A.4 Comparison between fitness centrality and average nearest neighbors degree

Figure 17 illustrates the relationship between the average nearest neighbors degree (Knn) and fitness centrality for different real-world networks. The scatter plot depicts the dependence of these two metrics. Moreover, in Table 6, we provide the least square estimate of a possible power-law relationship. In the cases of ZKC and Elegans, there is a mild anti-correlation, with $R^2 = 0.67$ and 0.47 respectively. In contrast, the remaining three cases exhibit low R^2 values, indicating that the power-law relationship between Knn and fitness centrality is less pronounced in these instances. The Spearman rank correlation reported in the last column of Table 6 aligns with the previous consideration.

Table 6: Power-law dependence between Knn and fitness centrality. For the scatter plot in Fig. 17, we assume a power-law relationship between the quantities shown in the leftmost column and fitness. We estimate the exponent of this relationship using least square regressions. We report the Spearman rank correlation in the last column.

Network	Exponent	R2	Spearman Corr.
Zachary	-0.1	0.67	-0.88
Train	-0.04	0.06	-0.24
Jazz	0.1	0.08	0.09
C. elegans	-0.17	0.47	-0.57
Dolphins	-0.05	0.21	-0.46

A.5 Attack vulnerability of networks

In this Section, we provide additional details regarding the vulnerability analysis on the network. The attack procedure works as follows: a) a node metric, which can be the degree, betweenness, or fitness, is computed on the network; b) the node with the highest score is removed; c) the process is iterated until all the nodes are disconnected. This is equivalent to the "recalculated" variant (R) presented in (Holme et al., 2002), which typically performs better than the "initial" variant (I). In the I-variant, the metrics are not recomputed after a node is removed from the network, so that the order of the node removal is determined once at the beginning.

To assess the vulnerability of the graph under attack, we use the size of the largest connected component S_c, which reveals how many nodes remain in the largest connected part of the network. Additionally, we use the inverse geodesic length, which is defined as:

$$\langle l^{-1} \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \tag{A.8}$$

where d_{ij} is the shortest path between node i and j. This quantity is the average value of the inverse of the distance between nodes in such a way that two disconnected nodes contribute to 0. A large value of $\langle l^{-1} \rangle$ reflects high functionality and navigability in the network, while it approaches 0 when all nodes are disconnected. In the case of a weighted

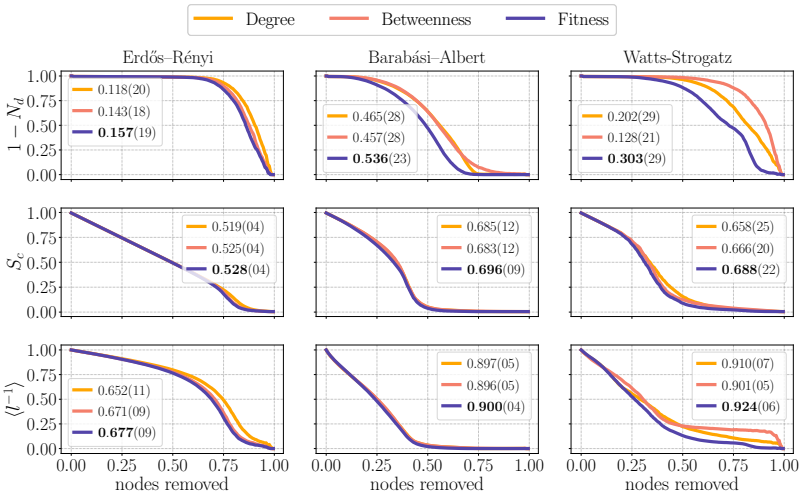

Figure 18: Attack vulnerability for synthetic networks (I-variant). Attacks based on fitness centrality consistently outperform or are comparable to those based on other measures. In particular, fitness centrality excels in terms of the number of disconnected nodes, N_d. The plots represent the average results over 100 realizations of random networks with 200 nodes, using the following parameters: ER: $p = 0.1$, BA: $m = 5$, WS: $k = 6$, $p = 0.1$. Values corresponding to the best performance (highest AOC) are highlighted in boldface.

network, the shortest path d_{ij} is calculated as the sum of the inverses of the weights of the links connecting nodes i and j.

Finally, we also considered the number of disconnected nodes N_d as a measure of the grade of network disruption under attack.

A.5.1 Initial attack (I-variant)

Figure 18 illustrates the behavior of $(1 - N_d)$, S_c, and $\langle l^{-1} \rangle$ as a function of the proportion of removed nodes in the I case for synthetic networks (Erdős–Rényi, Barabási-Albert, and Watts-Strogatz). With respect to all three quantities, the fitness strategy consistently results in the most rapid collapse of the entire network.

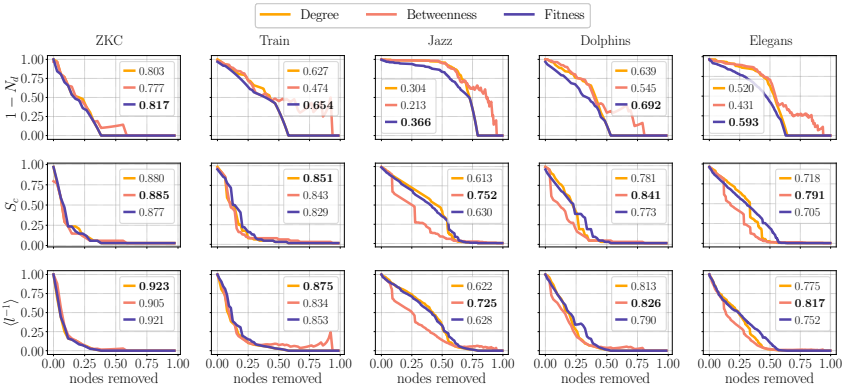

Figure 19: R-variant attack for real networks. In the R-variant, we recalculate the centrality measures after each node removal. The real networks are the same as those considered in the main text. Values corresponding to the best performance (highest AOC) are highlighted in boldface.

A.5.2 Recalculated attack (R-variant)

Figure 19 and Figure 20 illustrate the behavior of $(1 - N_d)$, S_c, and $\langle l^{-1} \rangle$ as a function of the proportion of removed nodes in the R case for real-world and synthetic networks, respectively. In all networks, fitness consistently outperforms the other measures when dealing with $(1 - N_d)$. With respect to S_c, and $\langle l^{-1} \rangle$, betweenness is typically the most effective strategy. In the Train, ZKC, and Barabási-Albert networks, degree is the optimal strategy for $\langle l^{-1} \rangle$.

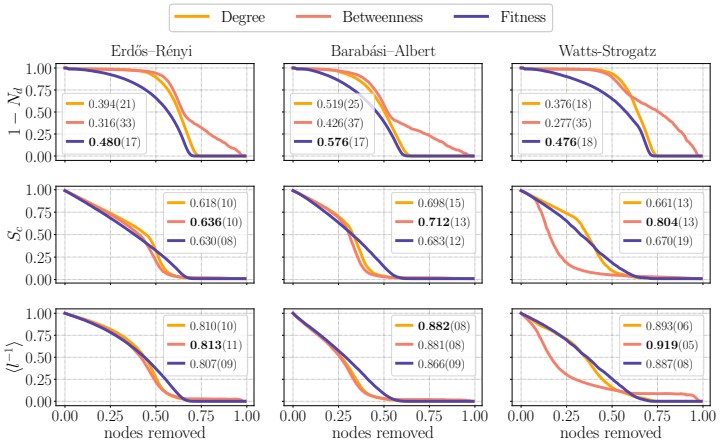

Figure 20: R-variant attack for synthetic networks. In the R-variant, we recalculate the centrality measures after each node removal. The plots represent the average results over 100 realizations of random networks with 200 nodes, using the following parameters: ER: $p = 0.1$, BA: $m = 5$, WS: $k = 6$, $p = 0.1$. Values corresponding to the best performance (highest AOC) are highlighted in boldface.

A.6 Weighted Networks

Figure 21 demonstrates the application of fitness centrality to weighted networks. The calculation is straightforward, requiring only a simple substitution of the weighted adjacency matrix in place of the binary adjacency matrix. Importantly, the underlying algorithmic structure remains unchanged. The top panel presents a scatter plot of strength, eigenvector, and betweenness centralities versus the logarithm of fitness. The data indicate a lack of correlation between fitness and eigenvector centrality, while we observe little correlation between fitness and strengths and between fitness and betweenness centrality. The bottom panel presents the results of vulnerability analysis (refer to Section A.5 for further details) in the I case for both the binary and the weighted versions. In both networks we obtain very similar results, with fitness consistently outperforming the other measures.

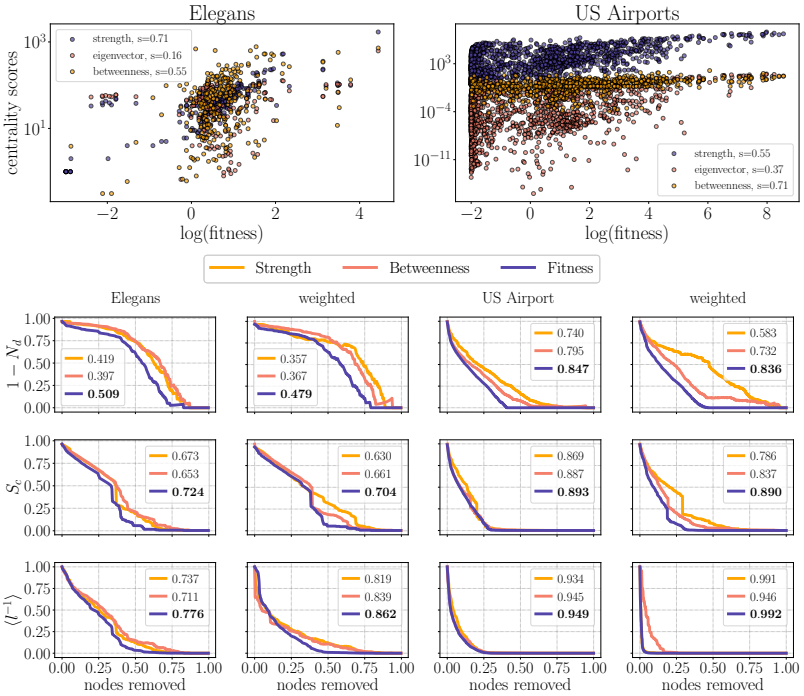

Figure 21: Weighted networks. Top: Centrality measures vs. the logarithm of fitness for the two weighted networks analyzed. Fitness (x-axis) is compared with strength, eigenvector, and betweenness centralities. s is the Spearman rank correlation. For the C. Elegans network, a significant correlation is observed only with strength centrality, while for the US Airports network, the strongest correlation is with betweenness centrality. In all other cases, correlations range from low to moderate. Note that the unweighted version of the C. Elegans network is discussed in the main text. **Bottom:** Attack vulnerability measures calculated in the I-case, i.e., without recalculating fitness after each node removal. The second and fourth columns show the results for the weighted networks, while the first and third columns present the unweighted versions of the networks, included for comparison. Definitions of the metrics $(1 - N_d)$, S_c, and $\langle l \rangle^{-1}$ are provided in the main text. The legends report the AOC score (Area Over the Curve), with the highest value highlighted in boldface. Fitness centrality consistently outperforms the other measures.

Appendix B

Supplementary Information Chapter 3

B.1 Data

In the following subsections, we describe the six food webs we analyzed. The edgelists and species' names for St.Marks, Coachella, and Benguela were constructed by examining the adjacency matrices in their original publications (Christian and Luczkovich, 1999; Polis, 1991; Yodzis, 1998). The edgelists and species' names for the Florida bay, Cypress dry season, Mangrove wet season (Ulanowicz, Bondavalli, and Egnotovich, 1998) were originally available at DuBois et al. (2003), sourced from Batagelj and Mrvar (2006). All the edgelists and species' names are available in the Zenodo repository (Calò, 2025b).

We removed three nodes (Input, Output, and Respiration) from the Florida Bay, Cypress Dry Season, and Mangrove Wet Season networks to maintain consistency across datasets. For the St. Marks, only step 3 (adding the links between the root node and all the other nodes) was performed, since the first two were already present in the data. Furthermore, we manually added one missing link, comparing the data with the adjacency matrix present in the original paper (Christian and Luczkovich, 1999). These preprocessing steps ensure that all networks represent com-

parable ecological entities.

B.1.1 Florida bay

This network represents trophic flows through the ecosystem of Florida Bay. It includes various species such as seagrass and algae producers, microfauna, macroinvertebrates, fishes, mammals, and avifauna (Ulanowicz, Bondavalli, and Egnotovich, 1998).

B.1.2 Cypress dry season

This network depicts the cypress wetlands in South Florida during the dry season (Ulanowicz, Bondavalli, and Egnotovich, 1998).

B.1.3 Mangrove wet season

This network illustrates the mangrove ecosystem in South Florida during the wet season (Ulanowicz, Bondavalli, and Egnotovich, 1998).

B.1.4 St. Marks

This network describes a winter's *Halodule wrightii* (shoal grass) community in Goose Creek Bay, St. Marks National Wildlife Refuge, Florida (Christian and Luczkovich, 1999).

B.1.5 Coachella

This network represents the sand community in the Coachella Valley desert. The biota includes species of vascular plants, vertebrates, arachnids, microorganisms, and insects (Polis, 1991).

B.1.6 Benguela

This network characterizes the Benguela Marine Ecosystem, including marine mammals and fisheries (Yodzis, 1998).

Table 7: Network statistics for analyzed food webs.

Food web	Nodes	Edges	Density (%)
St. Marks	49	275	11.5
Benguela	30	234	26.0
Coachella	30	294	32.7
Florida	126	2078	13.1
Cypress	69	634	13.3
Mangrove	95	1439	15.9

B.2 Network statistics

Table 7 shows the number of nodes (n), the number of links (m), and the network density (d) for each food web. n includes an additional root node for each network, m considers self-loops, and d is given by $d = \frac{m}{n^2}$ (directed networks with self-loops).

B.3 Correlation with eigenvector centralities

Fig. 22 reveals a strong correlation between the importance index and eigenvector-in centrality (Spearman correlation coefficient = 0.93), explaining their similar performance in extinction area analysis.

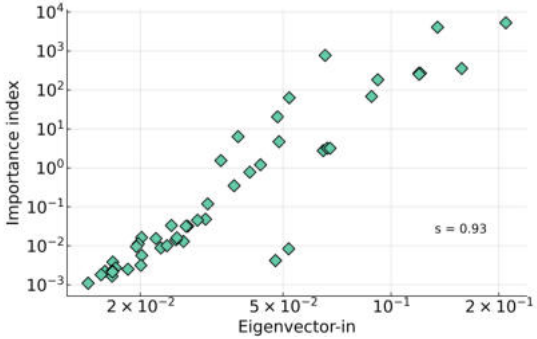

Figure 22: Relationship between eigenvector-in centrality and importance index in the Cypress food web. The Spearman corr. (s) is reported.

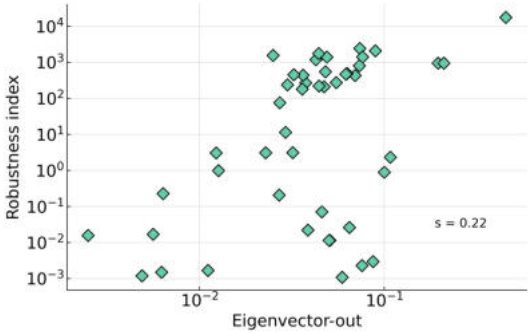

Figure 23: Relationship between eigenvector-out centrality and robustness index in the Cypress food web. The Spearman corr. (s) is reported.

Conversely, Fig. 23 shows a weak correlation between the robustness index and eigenvector-out centrality (Spearman correlation coefficient = 0.22), accounting for significant differences in identifying resilient species.

B.4 Cascading extinctions

B.4.1 Recalculated attack

Table 8 compares mean extinction areas using different ranking strategies, recomputing rankings after each species removal. The results reveal that in this case eigenvector-in centrality generally outperforms the importance index, which in turn outperforms in-degree. Notably, for the Coachella food web, the importance index shows superior performance compared to eigenvector-in centrality.

B.4.2 Recalculated vs. initial attack

Table 9 shows the maximum mean extinction areas for each food web and ranking method (in-degree, eigenvector-in, and importance index). The reported value represents the better result between the fixed and re-

Table 8: Comparison of mean extinction area across ecosystems using different species ranking methods. We compute the in-degree, eigenvector-in, and importance index rankings at the outset and recompute them after each step of the simulated extinction process. Values represent means and standard deviation. The standard deviation for the eigenvector-in and importance index is smaller than the last significative digit displayed (except for the eigenvector-in in the Coachella food web, for which it is 1.0). Higher percentages indicate greater ecosystem sensitivity to species loss based on the respective ranking method. Bold values indicate the best-performing method for each ecosystem.

Food Web	In-degree (%)	Eigenvector-in (%)	Importance index (%)
St. Marks	84.2 ± 1.1	**91.5**	87.4
Benguela	78.2 ± 0.3	**94.7**	87.9
Coachella	71.5 ± 0.1	71.6	**79.3**
Florida	63.5 ± 0.0	**87.3**	77.8
Cypress	78.6 ± 0.3	**86.6**	82.2
Mangrove	68.1 ± 0.4	**95.3**	92.3

Table 9: Comparison of maximum mean extinction areas across ecosystems. For each food web and ranking method (in-degree, eigenvector-in, and importance index), the table presents the higher mean extinction area obtained from either the fixed or recomputed ranking strategy.

Food Web	In-degree (%)	Eigenvector-in (%)	Importance index (%)
St. Marks	84.8 ± 1.1	**91.6**	91.4
Benguela	78.2 ± 0.3	**94.7**	91.3
Coachella	71.5 ± 0.1	75.9	**79.3**
Florida	64.6 ± 0.3	**87.3**	82.6
Cypress	78.6 ± 0.3	86.6	**86.9**
Mangrove	72.5 ± 0.1	**95.3**	94.3

computed ranking strategies, i.e., initial and recalculated attack variants.

In Fig. 24, we compare the extinction curves of all food webs under the two distinct attack strategies. In both scenarios, species are ranked using the importance index. We notice that Coachella is the only food web where recomputing outperforms the fixed ranking.

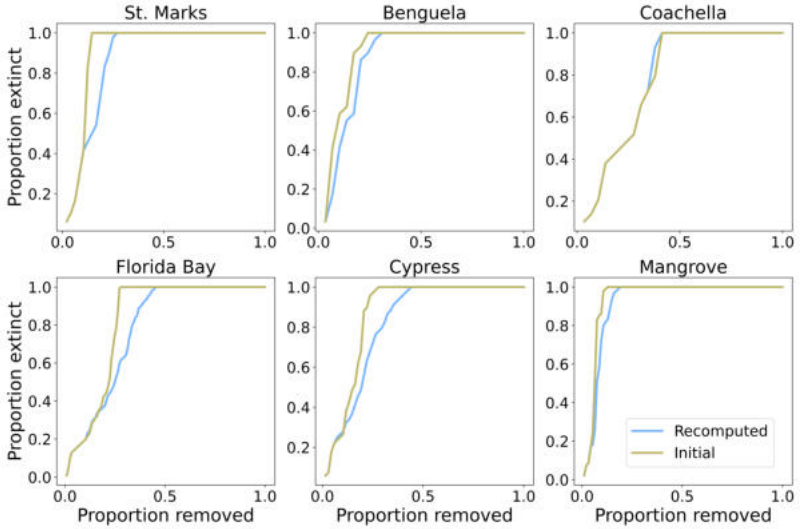

Figure 24: Comparative extinction curves across food webs. Extinction trajectories for all six food webs, illustrating species loss under two removal strategies: dynamically recomputed importance index ranking and fixed initial importance index ranking. Each subplot depicts the proportion of extinct species versus the proportion of species removed, highlighting differences between adaptive and static removal processes.

B.5 Robustness of species

Table 10 presents the Spearman rank correlation coefficients between removal steps and two ecosystem metrics: average log robustness index and average log eigenvector-out centrality of extinct species. These coefficients were calculated based on species removals following the importance index ranking. The results demonstrate that the robustness index consistently outperforms eigenvector-out centrality in identifying resilient species across all food webs.

This superior performance of the robustness index holds true even when species removals are based on eigenvector-in centrality, as illustrated in Fig. 25a for the Cypress food web.

Table 10: Spearman rank correlation coefficients (%) between removal steps and ecosystem metrics: average log robustness index and average log eigenvector-out of extinct species. Removals were performed following the importance index ranking.

Food Web	Robustness index (%)	Eigenvector-out (%)
St. Marks	100.00	92.86
Benguela	100.00	80.00
Coachella	100.00	60.71
Florida	91.93	55.09
Cypress	99.09	54.55
Mangrove	100.00	90.48

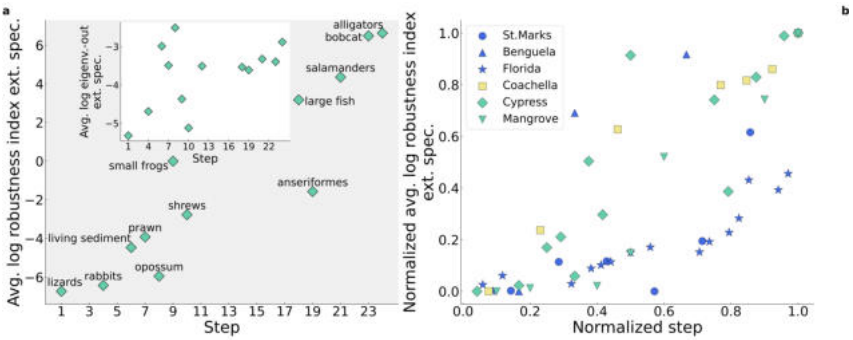

Figure 25: Ecological robustness: Species removal patterns in diverse food webs. Left: Cypress food web: Average log robustness index of extinct species vs. extinction step, based on eigenvector-in ranking. Labels indicate representative species becoming extinct at each step. **Inset:** Average log eigenvector-out centrality of extinct species vs. extinction step. **Right:** Normalized average log robustness index vs. normalized extinction step for multiple food webs. Values are log-transformed and scaled to [0,1] for cross-ecosystem comparison. This plot illustrates ecosystem resilience patterns and differential impacts of species loss across varied food web structures.

Fig. 25b extends this analysis to all six food webs, showcasing normalized average log robustness index versus normalized extinction steps. This comparison reveals consistent patterns of ecosystem resilience across diverse food web structures.

In the specific case of the Cypress food web, Fig. 26 provides addi-

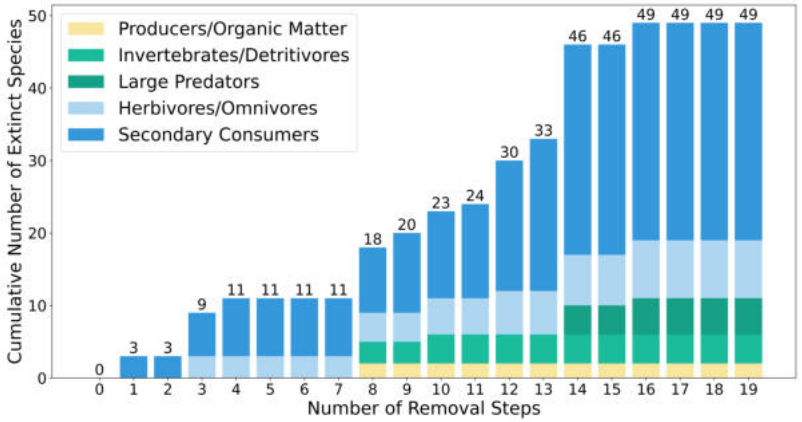

Figure 26: Cumulative species extinctions by functional category in the Cypress food web, based on the importance index ranking. Stacked bar chart illustrating the progressive accumulation of extinct species at each removal step. Colors represent different functional categories. Bar heights remain constant when no new extinctions occur, reflecting the cumulative nature of the data. This visualization demonstrates the differential impact of species removals on various functional groups within the ecosystem.

tional insights into the extinction dynamics. The visualization reveals that large predators, primarily Apex Trophic Nodes, persist until later removal stages, consistent with their high robustness index scores and broad prey bases.

B.6 Aggregation analysis

In the Florida Bay food web, we aggregated eight phytoplankton species (2μm Spherical Phytoplankt, Synedococcus, Oscillatoria, Small Diatoms ($< 20\mu$m), Big Diatoms ($> 20\mu$m), Dinoflagellates, Other Phytoplankton, Benthic Phytoplankton) into a single "Phytoplankton" node. We also aggregated two heron categories (Big Herons & Egrets, Small Herons & Egrets) into a single "Herons & Egrets" node, three turtle species (Loggerhead Turtle, Green Turtle, Hawksbill Turtle) into "Turtles", and three killifish species (Other Killifish, Goldspotted killifish, Rainwater killi-

fish) into "Killifish". We then examined the effect of these individual aggregations on the total extinction area resulting from simulated species removals and cascading effects. Aggregating the eight phytoplankton species led to the most substantial increase in the extinction area (eigenvector centrality-based area to 91.0, importance-based area to 90.1). The aggregation of Herons & Egrets resulted in a minor change. Aggregating turtles led to a slight decrease in the eigenvector centrality-based extinction area (to 82.1) and a minor increase in the importance-based area (to 82.4). Aggregating killifish resulted in a slight decrease in both the eigenvector centrality-based extinction area (to 82.3) and the importance-based area (to 82.1).

B.7 Cypress food web: Chat GPT categories

The following are the five categories into which the species of the Cypress food web have been divided, as generated by Chat GPT.

- **Producers/Organic Matter**: Living POC, Living sediment, Phytoplankton, Float. vegetation, Periphyton/Macroalgae, Macrophytes, Epiphytes, Understory, Vine Leaves, Hardwoods Leaves, Cypress Leaves, Cypress Wood, HW Wood, Roots
- **Invertebrates/Detritivores**: Crayfish, Apple Snail, Prawn, Aquatic Invertebrates, Ter. Invertebrates, Refractory Det., Liable Det., Vertebrate Det.
- **Herbivores/Omnivores**: Small Fish, herb and omniv, Small Frogs, Tadpoles, Galliformes, White ibis, Gruiformes, Caprimulgiformes, Hummingbirds, Woodpeckers, Passeriformes onniv., Squirrels, Mice & Rats, Rabbits, White-Tailed Deer, Armadillo
- **Secondary Consumers**: Small Fish, prim. carniv, Large Fish, Turtles, Lizards, Snakes, Salamanders, Large Frogs, Medium Frogs, Salamander L, Pelecaniformes, Anseriformes, Vultures, Egrets, Great blue heron, Other herons, Wood stork, Passeriformes pred., Opossum, Shrews, Bats, Mink, Otter, Bobcat

- **Large Predators**: Alligators, Kites & Hawks, Owls, Black Bear, G. Fox, Raccoon, Florida Panther, Hogs

B.8 Species categorization in robustness index-importance index plane

Species in each food web are categorized into four quadrants based on their robustness index and importance index: Base Trophic Nodes, Sheltered Trophic Nodes, Apex Trophic Nodes, and Carbon Bridge Nodes. Table 11 presents the species categorization for each food web based on their position in the robustness index-importance index plane. The Carbon Bridge Nodes for each relevant food web are as follows:

- Florida Bay: Benthic POC and Water POC.
- Cypress: Terrestrial Invertebrates, Liable Detritus, Refractory Detritus, and Vertebrate Detritus.
- Mangrove: POC, INSCT, and C in SED.

Table 11: Species categorization based on their trophic roles.

Food Web	Base Trophic Nodes	Sheltered Trophic Nodes	Apex Trophic Nodes
St. Marks	Benthic bacteria, Microfauna, Meiofauna, Bacterioplankton, Microprotozoa, Epiphyte-grazing amphipods, Suspension-feeding molluscs, Deposit-feeding peracaridan crustacean, Deposit-feeding gastropods, Other gastropods, Deposit-feeding polychaetes, Zooplankton, Halodule, Micro-epiphytes, Macro-epiphytes, Benthic algae, Phytoplankton, Detritus	Hermit crabs, Spider crabs (herbivores), Isopods, Brittle stars, Herbivorous shrimps, Tonguefish, Sheepshead minnow, Epiphyte-grazing gastropods, Suspension-feeding polychaetes	Omnivorous crabs, Blue crabs, Predatory shrimps, Catfish and stingrays, Gulf flounder and needlefish, Southern hake and sea robins, Atlantic silverside and bay anchovies, Killifishes, Gobies and blennies, Pipefish and seahorses, Red drum, Predatory gastropods, Predatory polychaetes, Benthos-eating birds, Fish-eating birds, Fish and crustacean-eating birds, Gulls, Raptors, Herbivorous ducks
Benguela	Phytoplankton, Benthic filter-feeders, Bacteria, Microzooplankton, Mesozooplankton, Macrozooplankton	Benthic carnivores, Gelatinous zooplankton, Anchovy, Pilchard, Round herring, Lightfish, Lanternfish, Goby	Other pelagics, Horse mackerel, Chub mackerel, Other groundfish, Hakes, Squid, Tunas, Snoek, Kob, Yellowtail, Geelbek, Whales & dolphins, Birds, Seals, Sharks
Coachella	plants/plant products, detritus, carrion, soil microbes	soil microarthropods and nematodes, surface arthropod detritivores, surface arthropod herbivores, herbivorous mammals and reptiles	soil micropredators, soil macroarthropods, soil macroarthropod predators, small arthropod predators, medium arthropod predators, facultative arthropod predators, life-history arthropod omnivore, spider parasitoids, primary parasitoids, hyperparasitoids, facultative hyperparasitoids, primarily herbivorous mammals and birds, small omnivorous mammals and birds, predaceous mammals and birds, arthropodivorous snakes, primarily arthropodivorous lizards, primarily carnivorous lizards, primarily carnivorous snakes, large primarily predaceous birds, large primarily predaceous mammals, golden eagle
Florida Bay	2μm Spherical Phytoplankt, Synedococcus, Oscillatoria, Small Diatoms ($< 20\mu$m), Big Diatoms ($> 20\mu$m), Dinoflagellates, Other Phytoplankton, Benthic Phytoplankton, Thalassia, Halodule, Syringodium, Drift Algae, Epiphytes, Water Flagellates, Water Ciliates, Acartia Tonsa, Oithona nana, Paracalanus, Other Copepoda, Meroplankton, Other Zooplankton, Benthic Flagellates, Benthic Ciliates, Meiofauna, Sponges, Coral, Other Cnidaridae, Echinoderma, Bivalves, Detritivorous Gastropods, Epiphytic Gastropods, Predatory Gastropods, Detritivorous Polychaetes, Predatory Polychaetes, Suspension Feeding Polych, Macrobenthos, Benthic Crustaceans, Detritivorous Amphipods, Herbivorous Amphipods, Isopods, Herbivorous Shrimp, Predatory Shrimp, Pink Shrimp, Thor Floridanus, Lobster, Detritivorous Crabs, Omnivorous Crabs, Predatory Crabs, Callinectus sapidus, Stone Crab, Sardines, Anchovy, Bay Anchovy, Lizardfish, Catfish, Eels, Toadfish, Brotalus, Halfbeaks, Needlefish, Other Killifish, Goldspotted killifish, Rainwater killifish, Sailfin Molly, Silverside, Other Horsefish, Gulf Pipefish, Dwarf Seahorse, Other Snapper, Mojarra, Grunt, Pinfish, Scianids, Parrotfish, Mullet, Blennies, Code Goby, Clown Goby, Flatfish, Filefishes, Puffer, Other Pelagic Fishes, Other Demersal Fishes	Roots, Sharks, Rays, Bonefish, Snook, Grouper, Jacks, Pompano, Gray Snapper, Porgy, Spotted Seatrout, Red Drum, Spadefish, Mackerel, Big Herons & Egrets, Small Herons & Egrets, Ibis, Roseate Spoonbill, Herbivorous Ducks, Omnivorous Ducks, Gruiformes, Small Shorebirds, Gulls & Terns, Kingfisher, Loggerhead Turtle, Green Turtle, Hawksbill Turtle, Manatee, DOC	Tarpon, Barracuda, Loon, Greeb, Pelican, Comorant, Predatory Ducks, Raptors, Crocodiles, Dolphin
Cypress	Living POC, Living sediment, Phytoplankton, Float. vegetation, Periphyton/Macroalgae, Macrophytes, Epiphytes, Understory, Hardwoods Leaves, Crayfish, Apple Snail, Prawn, Aquatic Invertebrates, Small Fish herb and omniv, Passeriformes onniv.	Vine Leaves, Cypress Leaves, Cypress Wood, HW Wood, Roots, Lizards, Medium Frogs, Small Frogs, Salamander L, Tadpoles, Anseriformes, Vultures, Galliformes, Caprimulgiformes, Hummingbirds, Woodpeckers, Passeriformes pred., Opossum, Shrews, Bats, Rabbits, White-Tailed Deer	Small Fish prim. carniv, Large Fish, Alligators, Turtles, Snakes, Salamanders, Large Frogs, Pelecaniformes, Kites & Hawks, Egrets, Great blue heron, Other herons, Wood stork, White ibis, Gruiformes, Owls, Black Bear, G. Fox, Raccoon, Mink, Otter, Florida Panther, Bobcat, Squirrels, Hogs, Armadillo
Mangrove	PHY, OTH. PP, LEAF, MICR. HO, ZOOPL., BACT.SED., FLA. SED., CIL. SED., MEIOF., MERO, EPIFN, POLY, AGAST, BVLVS, MBENTH, SCRUST, AMPHI, PENAID, CARID, OSHMP, JLOBST, OCRAB, TCRAB, DCRAB, PCRAB, FWINV, LARV, SPIDR, HERR, ANCH, KILLI, POEC, HRSE, SLVR, FWFSH, MOJA, GOBY, DOC	WOOD, ROOT, TGAST, RAYS, EFISH, PIN, MULL, LZRD, VULT, C & C, WOODP, PASSOMN, PASSPERD, RABT, SQIUR, M & R, DERS, MANA	SHRK, TARP, NEED, SNOOK, CUDA, SNAP, SCIAE, OFISH, TURT, SNKS, COCO, AMPH, L & G, PELC, CORM, BH & E, SE & E, IBIS, DUCK, DUCK, DUCK, K & H, MRAPT, GUIF, SSBIRDS, G & T, OWLS, OPSU, FOX, BEAR, RACO, M & O, CATS, DOLP

Appendix C

Supplementary Information Chapter 4

C.1 Data

The region-to-region OD matrices are publicly available (*ESPON Database Portal* 2020). Moreover, data and code are available in the Zenodo repository Calò, 2025a.

The data encompass the flows between 297 European regions recorded annually. Different periods are covered for each flow type, i.e., 2010–2014 for Erasmus, 2015–2020 for Horizon 2020, and 2010–2018 for all other categories. Distinct methodologies were employed by the researchers who collected and processed each type of OD dataset. They gathered and harmonized various data sources at both European and national levels, initially focusing on country-to-country flows. While most of the data were raw, some flows were estimated using specific techniques (Detailed information on the data sources and processing methodologies used for harmonization and estimation of missing values can be found in *IRiE* 2022). These country-level flows were then decomposed to the regional level for more detailed analysis. In Table 12, we present an overview of the data used in this study. It is important to note that the column "Methodology" describes the procedures employed by the original data

collectors, while the column "Our Analysis" outlines the additional steps we performed for our specific analysis.

Table 12: Flow type overview, data sources, methodologies, our analysis.

Flow Type	Description	Sources	Methodology	Our Analysis
Migration	Number of people migrating between NUTS 2 regions	EUROSTAT and NSI	Multi-step process: Base Data, Stock Gain estimation, In-Out-Cross estimation. Country-to-country matrices created, gaps filled using stock-gain method and linear models. Region-to-region flows estimated by decomposing country-level data	Applied floor function to ensure whole numbers. Divided by origin region population (see SI)
Tourism	Number of tourists traveling between NUTS 2 regions	EUROSTAT and UNWTO for country-to-country; EUROSTAT for regional domestic arrivals	Completed gaps in country-to-country and disaggregated country-to-country to region-to-region. Methods include cross-referencing UNWTO indexes, interpolation/extrapolation, and gravity model analysis using GDP, arrivals, and distance data	Applied floor function. Divided by origin region population
FDI	Shareholders' funds (thousand euros) in foreign-owned companies	AMADEUS database (Bureau van Dijk)	Aggregated firm-level data. Included intraregional and interregional intra-national flows	Summed FDI across all sectors. Divided by origin region GDP (see SI)
Remittances	Regionalized bilateral remittance estimates (thousand euros)	EUROSTAT and World Bank	Estimated regional-level flows by regionalizing national-level data using ratio of regional to national migration flows	Divided by origin region GDP
Freight Transport	Total freight flow between NUTS 2 regions (thousand tons)	Various, for road, rail, maritime, and air transport	Performed consistency and plausibility checks. Developed disaggregation procedures where regionalized data unavailable	Divided by total outgoing flows from each region, multiplied by region's relative economic importance
Erasmus Student Mobility	Higher education student mobility between partner countries	European Commission datasets	Geocoded individual movements to NUTS 2 regions based on sending and receiving institutions	Divided by origin region population
Horizon 2020 Partnerships	Number of H2020 partnerships between NUTS 2 regions	CORDIS project and participant organization lists	Geocoded organizations to NUTS 2 regions. Counted partnerships with coordinating partners as senders and other partners as receivers.	No additional processing
Passenger Transport	Passenger flows between NUTS 2 regions for air, maritime, and rail transport	Various Eurostat datasets	Implemented appropriate disaggregation procedures where regionalized data unavailable	Summed air, maritime, and rail ($\times 1000$) passenger flows. Divided by origin region population

C.1.1 Population

Total population on January 1st, sourced from *ESPON Database Portal* 2020. Manual additions were made for Mariotte (France) and Ireland for specific years. We normalize the migration flows by dividing them by the population of the origin region to address a potential bias. According to the gravity model in migration studies, regions with larger populations tend to have higher absolute numbers of incoming and outgoing migrants. By normalizing the flows, we eliminate this population size bias, allowing for a more accurate comparison of regions' relative contributions to migration patterns. This approach is also applied to other types of flows in our study, ensuring consistent analysis across different flow categories.

C.1.2 GDP

Gross Domestic Product at current market prices (million euros), sourced from *ESPON Database Portal* 2020. We used 2011 values for regions lacking 2010 data. Manual additions were made for Iceland and Liechtenstein.

C.1.3 Single-Layer Network Construction

As a first step, we represent each of the eight flow types as a separate weighted directed network (a "layer"). In each layer, nodes correspond to regions, links represent flows between regions within a given year, and link weights denote the magnitude of these flows.

C.1.4 Multilayer Network Construction

To analyze the complex interactions between different types of flows, we construct a multilayer network for each year by integrating all single-flow layers. Since the same set of regions is present across all layers, with connections existing only within each layer (i.e., no direct links between a tourism node and an FDI node), the resulting structure is a multiplex network (Bianconi, 2018). Figure 27 provides a simplified, conceptual

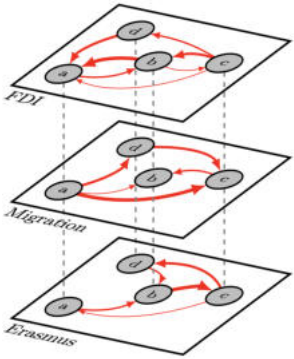

Figure 27: Illustrative schematic of a multiplex network. This conceptual diagram shows a network with three layers to represent different flow types. The nodes represent four hypothetical regions (a, b, c, and d), which are the same across all layers. The links are also hypothetical and are arranged differently on each layer for illustrative purposes, demonstrating how a region's connectivity can vary across dimensions.

illustration of this multiplex structure. The diagram uses three layers for visual clarity and employs generic link patterns to demonstrate the concept; it is not a direct representation of our empirical data, which encompass all available flow types.

C.2 Network properties

C.2.1 Network statistics

Table 13 presents the network statistics across European NUTS 2 regions from 2010 to 2018. For each flow type, we report the number of nodes, which indicates the participating NUTS 2 regions in the network, as well as the number of edges, representing the connections between these regions, where a connection signifies a non-zero flow. Additionally, we calculate the density of each network as the ratio of actual connections to the total possible connections, providing insight into how interconnected the regions are within each flow type.

Table 13: Network statistics for all layers.

Table 14: Migration, Tourism, Freight, and Erasmus.

Layer	Year	Nodes	Edges	Density(%)
Migration	2010	297	70797	80.53
Migration	2011	297	71535	81.37
Migration	2012	297	72459	82.42
Migration	2013	297	74502	84.75
Migration	2014	297	75342	85.70
Migration	2015	297	75580	85.97
Migration	2016	297	75955	86.40
Migration	2017	297	76044	86.50
Migration	2018	297	76281	86.77
Tourism	2010	297	87635	99.68
Tourism	2011	297	87649	99.70
Tourism	2012	297	87652	99.70
Tourism	2013	297	87665	99.72
Tourism	2014	297	87662	99.72
Tourism	2015	297	87669	99.72
Tourism	2016	297	87686	99.74
Tourism	2017	297	87682	99.74
Tourism	2018	297	87692	99.75
Freight	2010	297	46695	53.12
Freight	2011	297	47315	53.82
Freight	2012	297	46820	53.26
Freight	2013	297	46663	53.08
Freight	2014	297	46703	53.12
Freight	2015	297	44742	50.89
Freight	2016	297	43628	49.63
Freight	2017	297	43326	49.28
Freight	2018	297	42587	48.44
Erasmus	2010	265	21551	30.80
Erasmus	2011	264	22094	31.82
Erasmus	2012	270	23405	32.22
Erasmus	2013	272	24027	32.60
Erasmus	2014	274	24576	32.85

Table 15: FDI, Passengers, Remittances, and Horizon2020.

Layer	Year	Nodes	Edges	Density(%)
FDI	2010	292	24872	29.27
FDI	2011	292	25028	29.45
FDI	2012	292	25178	29.63
FDI	2013	292	25309	29.79
FDI	2014	292	25360	29.85
FDI	2015	292	25422	29.92
FDI	2016	292	25483	29.99
FDI	2017	292	25521	30.03
FDI	2018	292	25441	29.94
Passengers	2010	297	12144	13.81
Passengers	2011	297	12486	14.20
Passengers	2012	297	12547	14.27
Passengers	2013	297	12604	14.34
Passengers	2014	297	12649	14.39
Passengers	2015	297	12588	14.32
Passengers	2016	297	12716	14.46
Passengers	2017	297	12849	14.62
Passengers	2018	297	12998	14.79
Remittances	2010	297	81375	92.56
Remittances	2011	297	81355	92.54
Remittances	2012	297	81410	92.60
Remittances	2013	297	81243	92.41
Remittances	2014	297	81172	92.33
Remittances	2015	297	81016	92.16
Remittances	2016	297	81236	92.41
Remittances	2017	297	81250	92.42
Remittances	2018	297	81250	92.42
Horizon2020	2015	287	7482	9.12
Horizon2020	2016	284	7135	8.88
Horizon2020	2017	288	6899	8.35
Horizon2020	2018	279	6475	8.35

C.2.2 First order properties

Fig. 28 presents the CCDF for all flow types in 2018. This plot closely resembles the CCDF for 2010, indicating a consistent stability in the strength distributions over time. The tail behavior across both years suggests that the potential heavy-tailed relationships and distribution characteristics for different flow types remain largely unchanged. This similarity underscores the persistent nature of the network structure and flow patterns in the European regional system, with minimal alterations in the relative strengths of connections across various domains between 2010 and 2018.

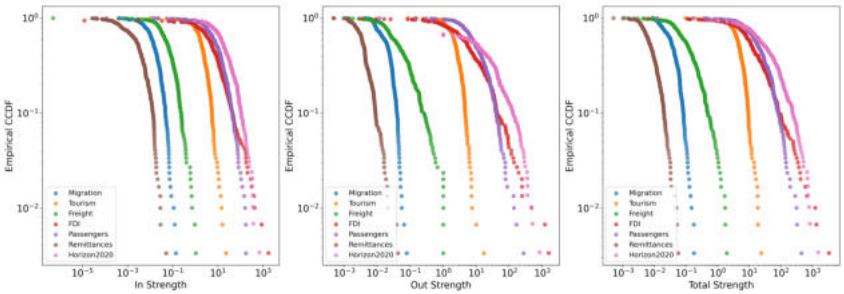

Figure 28: CCDF for the year 2018.

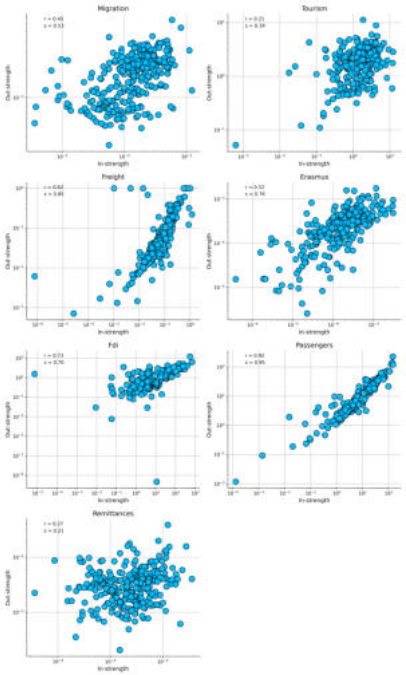

Figure 29: In- vs. out-strength (2010).

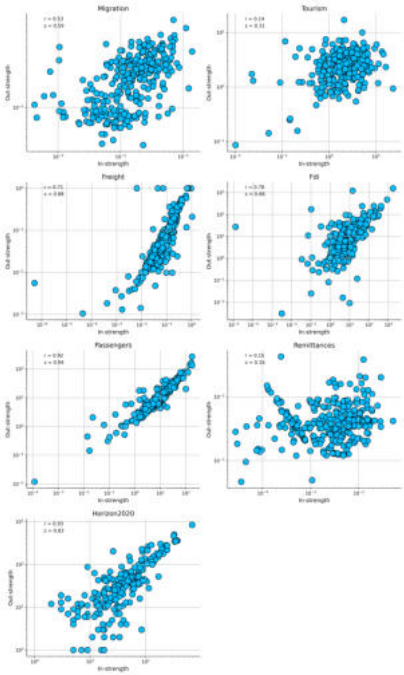

Figure 30: In- vs. out-strength (2018).

Figure 31: Comparison of in-strength vs. out-strength for 2010 and 2018.

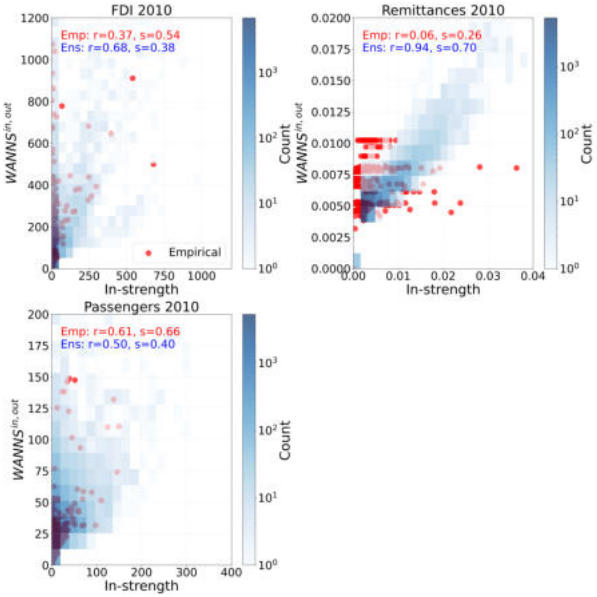

Figure 32: Comparison of empirical and ensemble WANNS for FDI, Remittances and Passengers in 2010. Each subplot shows the relationship between in-strength and WANNS^{in,out} values. The red points indicate empirical data, while the blue histogram represents the distribution of ensemble results. Correlation coefficients (Pearson's r and Spearman's s) are displayed for both empirical and ensemble data.

Fig. 31 depicts the relationship between in- and out-strength for all flow types in 2010 and 2018. The scatter plots reveal notable differences in correlations between various flow types, which remain consistent across both years. E.g., the Spearman correlation coefficients in 2010 range widely from 0.16 to 0.94. This substantial variation in correlations persists in 2018, suggesting that the underlying structures of these regional flow networks maintain their distinct characteristics over time.

C.2.3 Second order properties

Fig. 32 compares the empirical WANNS^{in,out} with 50 null model realizations for 2010 flows (FDI, Remittances, and Passengers).

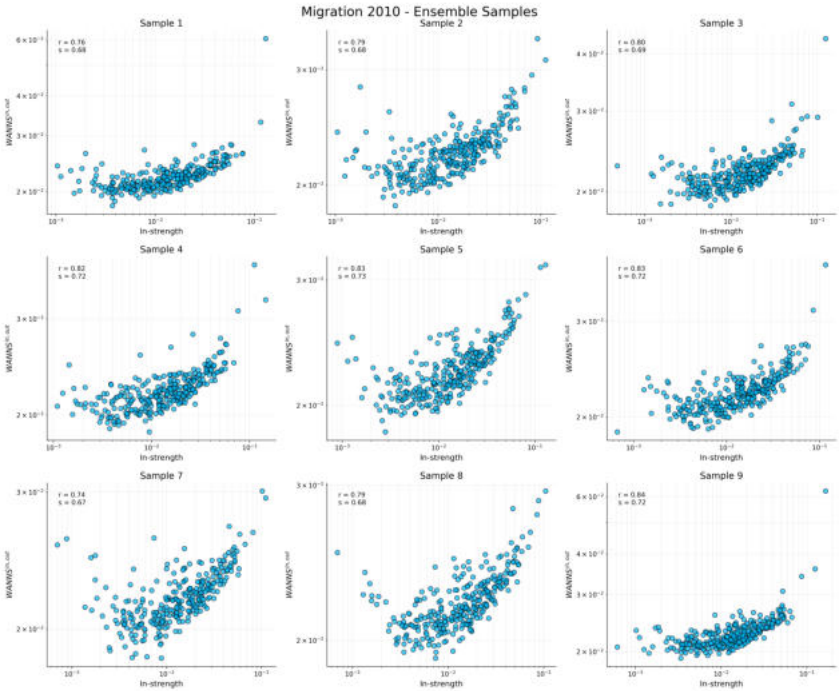

Figure 33: WANNS^{in,out} VS in-strength for ensemble copies for Migration the year 2010.

Fig. 33 focuses specifically on the Migration flow type in 2010, showing the relationship between the WANNS^{in,out} and the in-strength for ensemble copies. This visualization helps to understand the assortativity patterns in the migration network, revealing how regions with higher in-strength tend to receive flows from regions with higher out-strength.

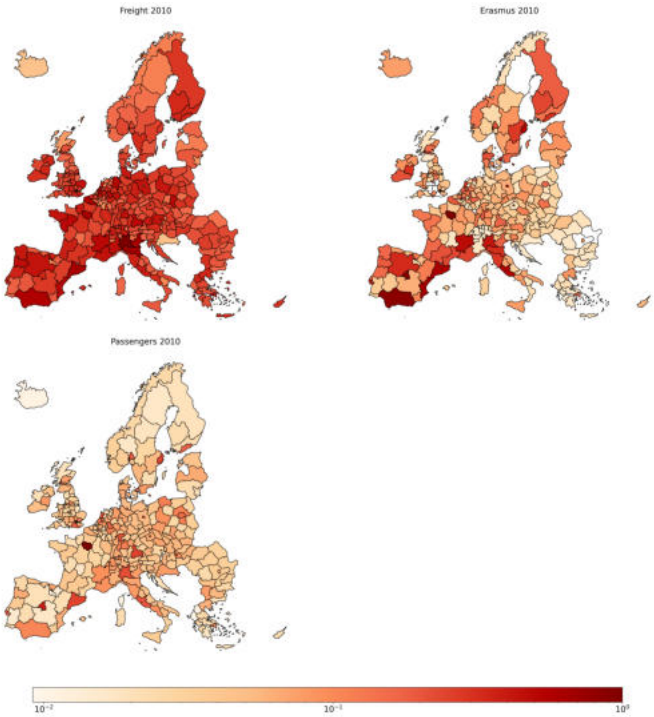

Figure 34: PageRank for Freight, Erasmus, and Passengers in 2010. Colors are displayed on a logarithmic scale, with values normalized such that the region with the highest centrality is set to 1.

C.3 PageRank

Fig. 34 illustrates the spatial distribution of PageRank centrality values across European regions for Freight, Erasmus, and Passengers in 2010.

C.3.1 Top 10 single-layer PageRanks

Tables 16 through 22 present the top 10 regions ranked by PageRank for various flow types in 2010. For migration (Table 16), Oberbayern emerges as the most central region, followed by Ile-de-France and In-

Table 16: Top 10 regions by PageRank for Migration in 2010.

Region	PageRank
Oberbayern	1.0000
Ile-de-France	0.8087
Inner London - East	0.6598
Darmstadt	0.6441
Stuttgart	0.6259
Düsseldorf	0.6252
Berlin	0.6167
Köln	0.6105
Attiki	0.6045
Lombardia	0.6024

Table 17: Top 10 regions by PageRank for Tourism in 2010.

Region	PageRank
Cataluña	1.0000
Ile-de-France	0.9793
Jadranska Hrvatska	0.9237
Andalucía	0.7983
Rhône-Alpes	0.6389
Provence-Alpes-Côte d'Azur	0.6367
Veneto	0.5567
Comunidad Valenciana	0.5176
Lombardia	0.4910
Comunidad de Madrid	0.4865

Table 18: Top 10 regions by PageRank for Freight in 2010.

Region	PageRank
Lombardia	1.0000
Zuid-Holland	0.7975
Emilia-Romagna	0.6890
Veneto	0.6428
Cataluña	0.6336
Ile-de-France	0.6021
Nord-Pas de Calais	0.5995
Andalucía	0.5737
Weser-Ems	0.5091
Comunidad Valenciana	0.5006

ner London - East. The list is dominated by German regions, highlighting their significance in migration patterns. Tourism (Table 17) shows Cataluña as the best region, with Ile-de-France and Jadranska Hrvatska following closely. This ranking reflects the popularity of Mediterranean coastal regions for tourism. In freight transport (Table 18), Lombardia leads, followed by Zuid-Holland and Emilia-Romagna, underscoring the importance of industrial and port regions in goods movement.

Table 19: Top 10 regions by PageRank for Erasmus in 2010.

Region	PageRank
Ile-de-France	1.0000
Comunidad de Madrid	0.9202
Andalucía	0.8484
Comunidad Valenciana	0.6656
Cataluña	0.6186
Berlin	0.5657
Lazio	0.5479
Stockholm	0.5269
Rhône-Alpes	0.4794
Hovedstaden	0.4272

Table 20: Top 10 regions by PageRank for FDI in 2010.

Region	PageRank
Noord-Holland	1.0000
Ile-de-France	0.6988
Comunidad de Madrid	0.5451
Lombardia	0.4559
Bucureşti - Ilfov	0.4537
Luxembourg	0.4333
Eastern and Midland	0.4220
Région de Bruxelles-Capitale	0.3791
Zuid-Holland	0.3750
Warszawski stołeczny	0.2924

Table 21: Top 10 regions by PageRank for Passengers in 2010.

Region	PageRank
Ile-de-France	1.0000
Comunidad de Madrid	0.3822
Inner London - West	0.3775
Inner London - East	0.3187
Stockholm	0.2466
Oberbayern	0.2449
Cataluña	0.2446
Berlin	0.2232
Noord-Holland	0.2197
Lombardia	0.2186

For Erasmus (Table 19), Ile-de-France ranks first, followed by Comunidad de Madrid and Andalucía, indicating the attractiveness of these regions for international students. In FDI (Table 20), Noord-Holland tops the list, with Ile-de-France and Comunidad de Madrid following, reflecting their financial importance. Passengers (Table 21) are dominated by Ile-de-France, with a major lead over Comunidad de Madrid and Inner London - West, highlighting Paris's role as a major transportation hub.

Table 22: Top 10 regions by PageRank for Remittances in 2010.

Region	PageRank
Ile-de-France	1.0000
Cataluña	0.9165
Luxembourg	0.7257
Comunidad de Madrid	0.7015
Vidurio ir vakaru Lietuvos regionas	0.6988
Prov. Antwerpen	0.5824
Rhône-Alpes	0.5389
Latvija	0.4810
Prov. Oost-Vlaanderen	0.4793
Comunidad Valenciana	0.4563

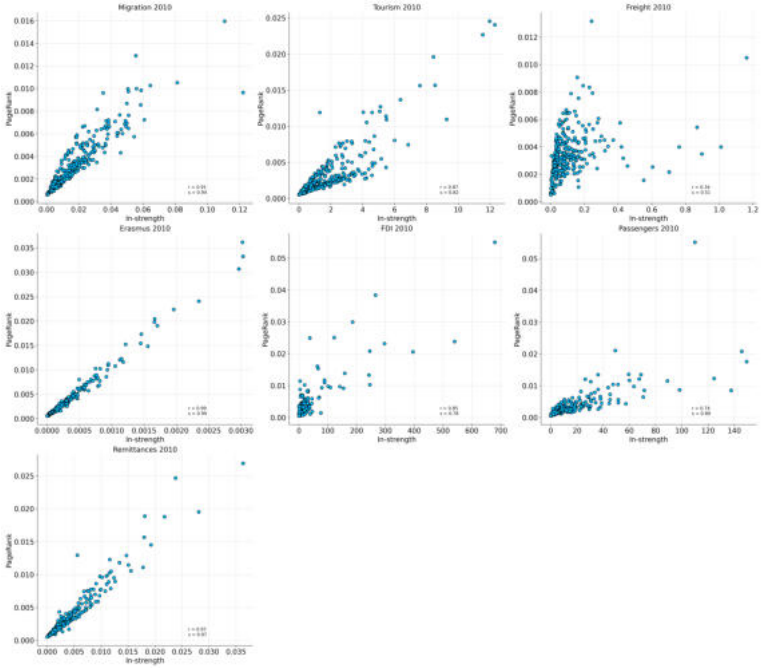

Figure 35: PageRank vs. in-strength for the year 2010.

Finally, for remittance flows (Table 22), Ile-de-France again leads, followed by Cataluña and Luxembourg, indicating the economic significance of these regions for international money transfers.

C.3.2 Single-layer PageRank vs. in-strength

Fig. 35 demonstrates the relationship between PageRank and in-strength across all flow types in 2010. The analysis reveals strong correlations for most flow types, with both Pearson and Spearman correlation coefficients exceeding 0.74. This indicates a robust association between a region's centrality and the volume of incoming flows for most networks. However, the Freight network stands out as an exception, exhibiting a notably weaker correlation.

C.3.3 Single-layer PageRank temporal evolution

To capture the temporal evolution of PageRank rankings, Figures 36–38 display a heatmap of the rankings for each layer, with regions ordered according to their average position across all layers and years.

Top and bottom migration trends

Fig. 39 illustrates the top 8 and bottom 8 slopes of PageRank trends for Migration, highlighting regions with notably increasing or decreasing centrality. Interestingly, London stands out among the top increasing trends, despite experiencing a noticeable dip in 2016, likely attributable to the Brexit referendum. This overall upward trajectory, even in the face of such a significant political event, underscores London's resilience and enduring importance as a migration hub.

C.3.4 Correaltion single-layer and multiplex PageRank

Fig. 40 displays the correlations between single-layer PageRank values and the multiplex PageRank for various flow types in 2010. The Spearman correlation coefficients range from 0.54 to 0.73, indicating moderate positive relationships between individual layer centralities and the overall multiplex centrality.

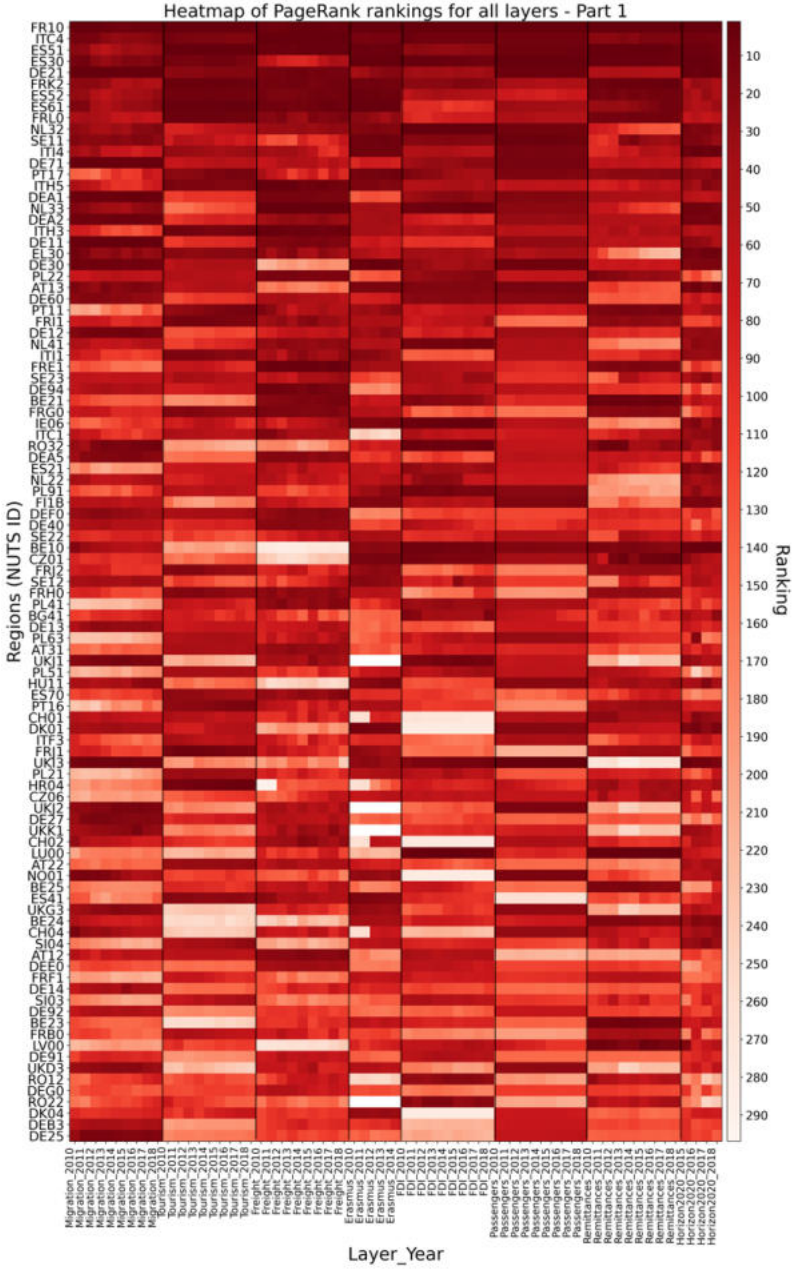

Figure 36: PageRank Heatmap - Top-Ranking regions across layers and years.

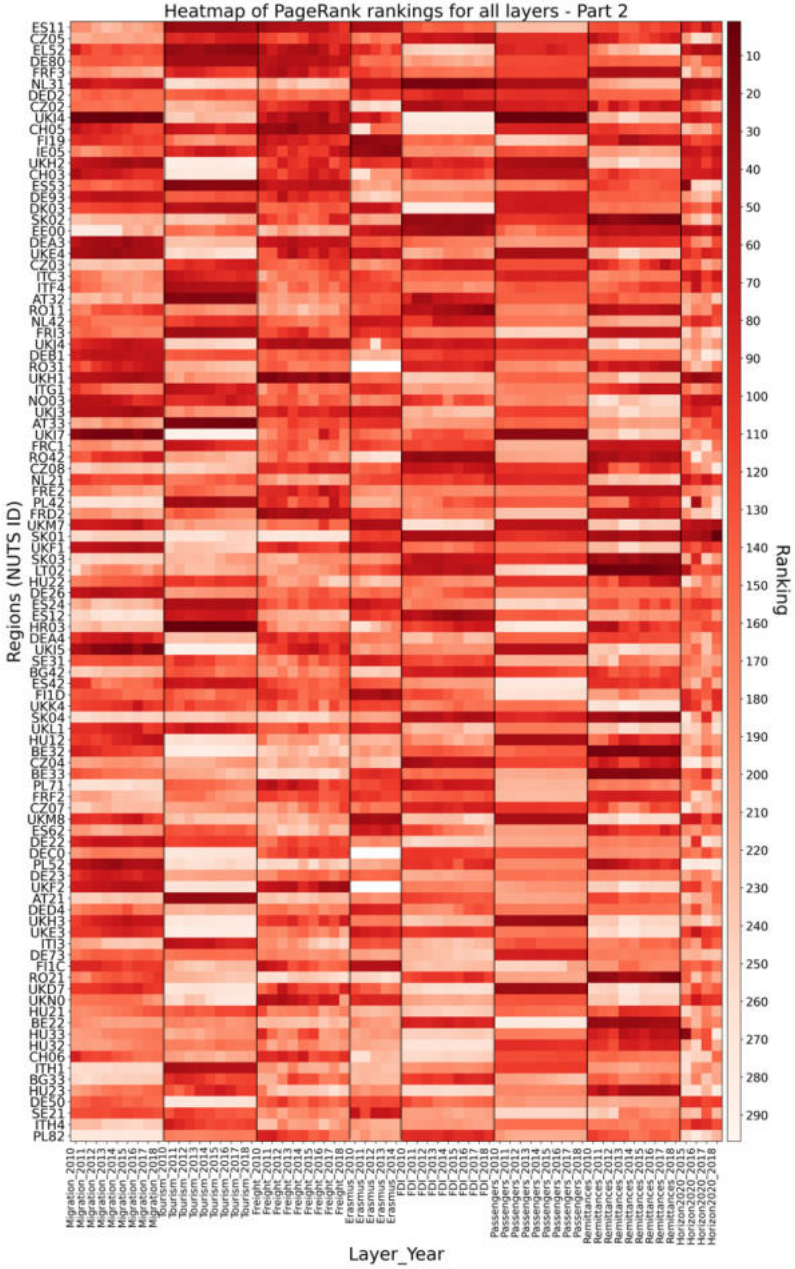

Figure 37: PageRank Heatmap - Mid-Ranking regions across layers and years.

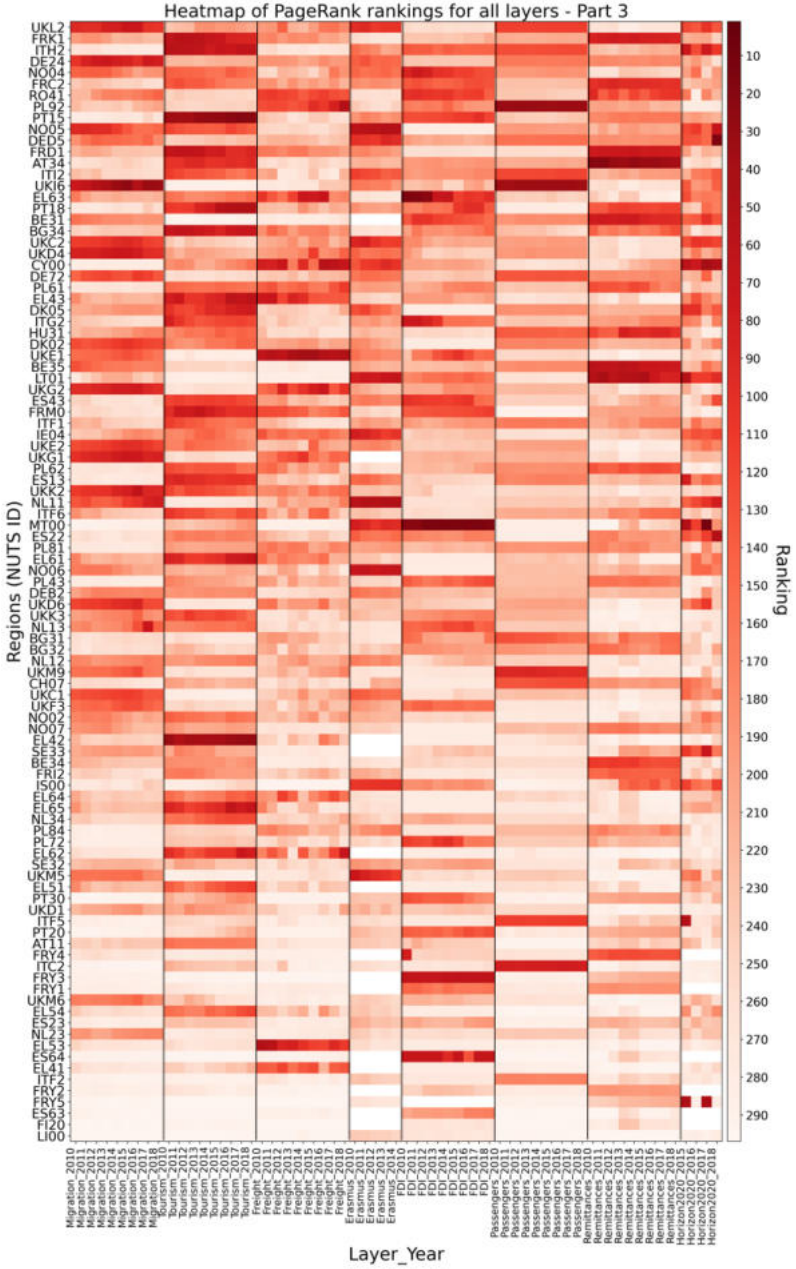

Figure 38: PageRank Heatmap - Bottom-Ranking regions across layers and years.

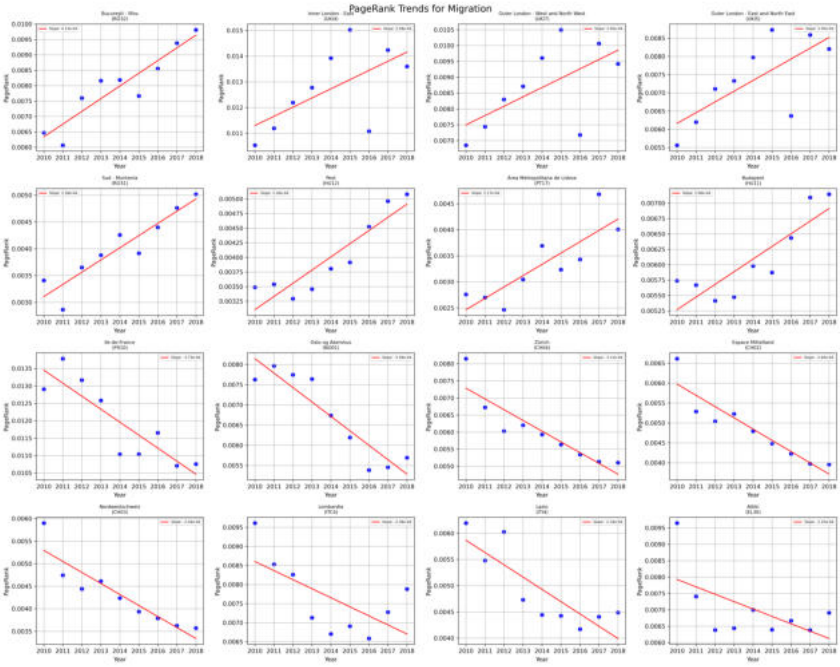

Figure 39: Top 8 and bottom 8 slopes of PageRank Trends for Migration.

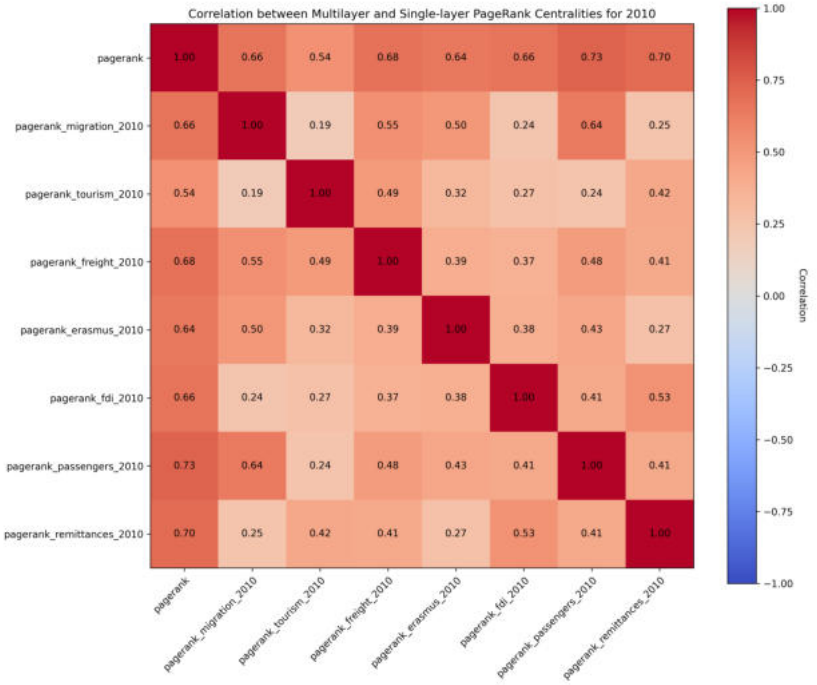

Figure 40: Correlations among single-layer PageRank and multiplex PageRank for 2010.

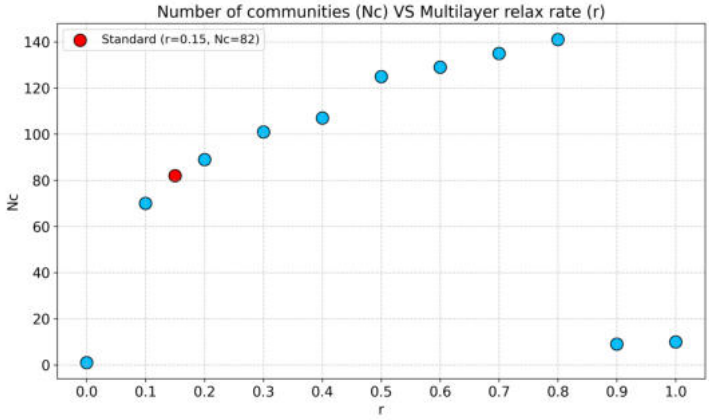

Figure 41: Number of communities VS multilayer relax rate for 2010.

C.4 Community detection

C.4.1 Infomap parameters

We configure the Infomap algorithm with the following parameters:

- `two_level=False`: Clusters the optimal number of nested modules, accommodating both country-level and single-region clusters.
- `num_trials=100`: Number of outer-most loops to run before selecting the best solution.
- `flow_model='rawdir'`: Determines node visitation rates based on the given direction and weight of edges, without using a PageRank algorithm.
- `entropy_corrected=True`: Corrects for negative entropy bias in small samples (many modules).
- `multilayer_relax_rate=0.15`: Probability of relaxing the constraint to move only within the current layer (default value).

In the Infomap community detection analysis of our multiplex network, it is possible for a region to be assigned to multiple communities across

different layers. To resolve such cases and provide a definitive community assignment, each region was ultimately assigned to the community in which it appeared most frequently across all layers.

C.4.2 Robustness check

Fig. 41 illustrates the relationship between the number of communities and the multilayer relax rate for 2010, showing that the range around the standard value (from 0.1 to 0.2) yields a relatively stable number of communities, varying from 70 to 89. Generally, increasing the relax rate r leads to a higher number of communities, until it reaches an extremely high value (0.9), at which point the number of communities sharply decreases to around 10.

C.4.3 List of detected communities

Table 23 presents the results of the Infomap community detection analysis.

Table 23: Infomap community detection results.

NUTS ID	Community	NUTS ID	Community	NUTS ID	Community	NUTS ID	Community	NUTS ID	Community	NUTS ID	Community
FRM0	1	BE23	4	CY00	8	HU23	14	DE11	23	DE93	40
FR10	1	BE22	4	UKK1	8	HU31	14	DE12	23	UKK4	41
FRL0	1	BE21	4	UKJ4	8	HU33	14	DE14	23	UKK3	41
FRE2	1	BE10	4	UKJ2	8	HU22	14	DE13	23	DE24	42
FRY3	1	NL34	4	UKJ3	8	HU32	14	UKN0	24	DE25	42
FRB0	1	BE25	4	RO11	9	HU12	14	IE04	24	DE94	43
FRK2	1	LU00	4	RO41	9	HU11	14	IE05	24	DED2	44
FRJ1	1	ITC2	5	RO21	9	HU21	14	IE06	24	DEG0	45
FRI3	1	ITC1	5	RO32	9	DK04	15	UKE2	25	FRF1	46
FRI2	1	ITH3	5	RO31	9	DK02	15	UKE3	25	FRY1	47
FRI1	1	ITC4	5	RO42	9	DK03	15	UKE4	25	FRY2	47
FRH0	1	ITI4	5	RO12	9	DK01	15	UKC2	25	DED4	48
FRJ2	1	ITF5	5	RO22	9	DK05	15	UKC1	25	FRF3	49
FRG0	1	ITF6	5	AT11	10	HR03	16	UKE1	25	ES62	50
ES70	2	ITF2	5	AT34	10	MT00	16	DE60	26	FRD2	51
ES64	2	ITI1	5	AT33	10	SI03	16	DE80	26	FRF2	52
ES63	2	ITH5	5	AT12	10	SI04	16	DEF0	26	ITF3	53
ES61	2	NL32	6	AT13	10	HR04	16	DE27	27	DE73	54
ES53	2	NL11	6	AT21	10	NO06	17	DE21	27	DEA4	55
ES51	2	NL12	6	AT22	10	NO02	17	UKD1	28	ITC3	56
ES42	2	NL13	6	AT32	10	NO03	17	UKD6	28	FRC1	57
ES23	2	NL21	6	AT31	10	NO05	17	UKD3	28	DEC0	58
ES24	2	NL33	6	SE23	11	NO07	17	UKD4	28	FRY4	59
ES30	2	NL42	6	SE33	11	NO04	17	UKD7	28	FRY5	59
ES41	2	NL31	6	SE21	11	NO01	17	UKM9	29	ITF4	60
ES11	2	NL23	6	SE22	11	EE00	18	UKM5	29	DE23	61
ES52	2	NL22	6	SE12	11	LV00	18	UKM6	29	DE72	62
PL62	3	NL41	6	SE11	11	LT02	18	UKM7	29	DE26	63
PL51	3	CZ06	7	SE32	11	LT01	18	UKM8	29	FRC2	64
PL52	3	SK04	7	FI20	11	CH04	19	DE30	30	DE22	65
PL61	3	CZ02	7	SE31	11	LI00	19	DE40	30	ITG1	66
PL63	3	CZ04	7	EL65	12	CH03	19	UKG3	31	ITI3	67
PL71	3	CZ03	7	EL53	12	CH05	19	UKG1	31	ITH2	68
PL72	3	CZ01	7	EL52	12	CH01	19	UKG2	31	ES22	69
PL81	3	CZ07	7	EL51	12	CH02	19	UKF1	32	ITI2	70
PL92	3	SK01	7	EL54	12	CH06	19	UKF2	32	DEB2	71
PL91	3	SK02	7	EL42	12	BG32	20	UKF3	32	ITH1	72
PL82	3	SK03	7	EL41	12	BG31	20	DE91	33	ITH4	73
PL43	3	CZ05	7	EL30	12	BG42	20	DE92	33	ITG2	74
PL41	3	CZ08	7	EL61	12	BG33	20	DE71	34	FRK1	75
PL42	3	UKH3	8	EL62	12	BG34	20	DEB1	35	PT15	76
PL84	3	UKJ1	8	EL63	12	BG41	20	DEB3	35	FRD1	77
PL21	3	UKK2	8	EL64	12	DEA2	21	DED5	36	ITF1	78
PL22	3	UKH1	8	EL43	12	DEA3	21	DEE0	36	ES43	79
BE35	4	UKH2	8	PT30	13	DEA5	21	ES12	37	ES13	80
BE34	4	UKI3	8	PT11	13	DEA1	21	ES21	37	CH07	81
BE33	4	UKI4	8	PT16	13	FI1D	22	UKL2	38	IS00	82
BE32	4	UKI5	8	PT17	13	FI1C	22	UKL1	38		
BE31	4	UKI6	8	PT18	13	FI1B	22	FRE1	39		
BE24	4	UKI7	8	PT20	13	FI19	22	DE50	40		

Bibliography

Albert, Réka and Albert-László Barabási (2002). "Statistical mechanics of complex networks". In: *Reviews of modern physics* 74.1, p. 47.

Allesina, Stefano, David Alonso, and Mercedes Pascual (2008). "A general model for food web structure". In: *Science* 320.5876, pp. 658–661.

Allesina, Stefano and Antonio Bodini (2004). "Who dominates whom in the ecosystem? Energy flow bottlenecks and cascading extinctions". In: *Journal of Theoretical Biology* 230.3, pp. 351–358.

Allesina, Stefano, Antonio Bodini, and Mercedes Pascual (2009). "Functional links and robustness in food webs". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1524, pp. 1701–1709.

Allesina, Stefano, Jacopo Grilli, et al. (2015). "Predicting the stability of large structured food webs". In: *Nature communications* 6.1, p. 7842.

Allesina, Stefano and Mercedes Pascual (2009). "Googling food webs: can an eigenvector measure species' importance for coextinctions?" In: *PLoS computational biology* 5.9, e1000494.

Amin, Ash (2004). "Regions unbound: towards a new politics of place". In: *Geografiska annaler: series B, human geography* 86.1, pp. 33–44.

Andrea Mazzolini Michele Caselle, Matteo Osella (2024). "Ranking nodes in bipartite systems with a non-linear iterative map". In: *arXiv preprint arXiv:2406.17572*.

Anselin, Luc (1988). *Spatial econometrics: methods and models*. Vol. 4. Springer Science & Business Media.

Aufiero, Sabrina et al. (2024). "Mapping job fitness and skill coherence into wages: an economic complexity analysis". In: *Scientific Reports* 14.1, p. 11752.

Balassa, Bela and Marcus Noland (1989). ""Revealed" Comparative Advantage in Japan and the United States". In: *Journal of International Economic Integration*, pp. 8–22.

Barabási, Albert-László and Réka Albert (1999). "Emergence of scaling in random networks". In: *science* 286.5439, pp. 509–512.
Barthélemy, Marc (2011). "Spatial networks". In: *Physics reports* 499.1-3, pp. 1–101.
Bartley, Timothy J et al. (2019). "Food web rewiring in a changing world". In: *Nature ecology & evolution* 3.3, pp. 345–354.
Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy (2009). "Gephi: An Open Source Software for Exploring and Manipulating Networks". In: *International AAAI Conference on Weblogs and Social Media*. URL: http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.
Batagelj, Vladimir and Andrej Mrvar (2006). *Pajek datasets*.
Bathelt, Harald and Johannes Glückler (2011). *The relational economy: Geographies of knowing and learning*. Oxford University Press.
Battiston, Stefano et al. (2012). "Debtrank: Too central to fail? financial networks, the fed and systemic risk". In: *Scientific reports* 2.1, pp. 1–6.
Benzi, Michele and Christine Klymko (2015). "On the limiting behavior of parameter-dependent network centrality measures". In: *SIAM Journal on Matrix Analysis and Applications* 36.2, pp. 686–706.
Bianconi, Ginestra (2018). *Multilayer networks: structure and function*. Oxford university press.
Blanchard, Julia L (2015). "A rewired food web". In: *Nature* 527.7577, pp. 173–174.
Boccaletti, Stefano et al. (2014). "The structure and dynamics of multilayer networks". In: *Physics reports* 544.1, pp. 1–122.
Bodini, Antonio, Giovanni Giavelli, and Orazio Rossi (1994). "The qualitative analysis of community food webs: implications for wildlife management and conservation". In: *Journal of environmental management* 41.1, pp. 49–65.
Bonaccorsi, Giovanni et al. (2019). "Country centrality in the international multiplex network". In: *Applied Network Science* 4.1, p. 126.
Brin, Sergey and Lawrence Page (1998). "The Anatomy of a Large-Scale Hypertextual Web Search Engine". In: *Computer Networks* 30, pp. 107–117.
Bryan, Kurt and Tanya Leise (2006). "The $25,000,000,000 eigenvector: The linear algebra behind Google". In: *SIAM review* 48.3, pp. 569–581.
Caenorhabditis elegans (neural) network dataset – KONECT (2018). Last accessed 06.2024. URL: http://konect.cc/networks/dimacs10-celegansneural.

Caldarelli, Guido (2007). *Scale-free networks: complex webs in nature and technology*. Oxford University Press.
Caldarelli, Guido et al. (2012). "A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy". In: *PLOS ONE* 7.10, pp. 1–11.
Calò, Emanuele (2025a). *European Region-to-Region Origin Destination Matrices*. Version 1. DOI: 10.5281/zenodo.15401969. URL: https://doi.org/10.5281/zenodo.15401969.
— (2025b). *Food webs: edge lists and species names*. DOI: 10.5281/zenodo.15045517. URL: https://doi.org/10.5281/zenodo.15045517.
Capocci, A. et al. (2005). "Detecting communities in large networks". In: *Physica A: Statistical Mechanics and its Applications* 352.2, pp. 669–676. ISSN: 0378-4371.
Castells, Manuel (1996). *The information age*. Vol. 1. s 81. Oxford Blackwell Publishers.
Cavallaro, Lucia et al. (2020). "Disrupting resilient criminal networks through data analysis: The case of Sicilian Mafia". In: *Plos one* 15.8, e0236476.
Christian, Robert R and Joseph J Luczkovich (1999). "Organizing and understanding a winter's seagrass foodweb network through effective trophic levels". In: *Ecological modelling* 117.1, pp. 99–124.
Cimini, Giulio et al. (2019). "The statistical physics of real-world networks". In: *Nature Reviews Physics* 1.1, pp. 58–71.
Cirtwill, Alyssa R et al. (2018). "A review of species role concepts in food webs". In: *Food Webs* 16, e00093.
Costantini, Lorenzo et al. (2022). "Measuring node centrality when local and global measures overlap". In: *Physical Review E* 105.4, p. 044317.
Cracolici, Maria Francesca and Peter Nijkamp (2009). "The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions". In: *Tourism management* 30.3, pp. 336–344.
Cristelli, Matthieu et al. (2013). "Measuring the intangibles: A metrics for the economic complexity of countries and products". In: *PloS one* 8.8, e70726.
Cudny, Waldemar and Josef Kunc (2022). *Growth and change in post-socialist cities of Central Europe*. Routledge, Taylor & Francis Group.
De Domenico, Manlio, Mason A Porter, and Alex Arenas (2015). "MuxViz: a tool for multilayer analysis and visualization of networks". In: *Journal of Complex Networks* 3.2, pp. 159–176.

De Domenico, Manlio, Albert Solé-Ribalta, et al. (2013). "Mathematical formulation of multilayer networks". In: *Physical Review X* 3.4, p. 041022.
De Marzo, Giordano and Vito D. P. Servedio (2023). "Quantifying the complexity and similarity of chess openings using online chess community data". In: *Scientific Reports* 13.1, p. 5327. ISSN: 2045-2322. URL: https://doi.org/10.1038/s41598-023-31658-w.
Diem, Christian et al. (2022). "Quantifying firm-level economic systemic risk from nation-wide supply networks". In: *Scientific reports* 12.1, p. 7719.
Docquier, Frédéric and Elisabetta Lodigiani (2010). "Skilled migration and business networks". In: *Open Economies Review* 21.4, pp. 565–588.
Dolphins network dataset – KONECT (2017). Last accessed 06.2024. URL: http://konect.cc/networks/dolphins.
Domínguez-García, Virginia and Miguel A. Muñoz (2015). "Ranking species in mutualistic networks". In: *Scientific Reports* 5.1, p. 8182. ISSN: 2045-2322. URL: https://doi.org/10.1038/srep08182.
DuBois, Christopher L. et al. (2003). *netdata: A Collection of Network Data*.
Dunhill, Alexander M et al. (2024). "Extinction cascades, community collapse, and recovery across a Mesozoic hyperthermal event". In: *Nature Communications* 15.1, p. 8599.
Dunne, Jennifer A (2006). "The network structure of food webs". In: *Ecological networks: linking structure to dynamics in food webs*, pp. 27–86.
Dunne, Jennifer A, Richard J Williams, and Neo D Martinez (2002). "Food-web structure and network theory: the role of connectance and size". In: *Proceedings of the National Academy of Sciences* 99.20, pp. 12917–12922.
Eklöf, Anna et al. (2013). "The dimensionality of ecological networks". In: *Ecology letters* 16.5, pp. 577–583.
ESPON Database Portal (2020). https://database.espon.eu/. Accessed: 2024-06-17.
Estrada, Ernesto (2007). "Characterization of topological keystone species: Local, global and "meso-scale" centralities in food webs". In: *Ecological Complexity* 4.1-2, pp. 48–57.
Fageda, Xavier (2017). "International air travel and FDI flows: Evidence from Barcelona". In: *Journal of Regional Science* 57.5, pp. 858–883.
Faludi, Andreas (2006). "From European spatial development to territorial cohesion policy". In: *Regional studies* 40.6, pp. 667–678.
Fussmann, Gregor F and Gerd Heber (2002). "Food web complexity and chaotic population dynamics". In: *Ecology Letters* 5.3, pp. 394–401.

Gabrielli, Andrea et al. (2019). “Grand canonical ensemble of weighted networks”. In: *Physical Review E* 99.3, p. 030301.
Garlaschelli, Diego, Guido Caldarelli, and Luciano Pietronero (2003). “Universal scaling relations in food webs”. In: *Nature* 423.6936, pp. 165–168.
Ghoshal, Gourab and Albert-László Barabási (2011). “Ranking stability and super-stable nodes in complex networks”. In: *Nature communications* 2.1, p. 394.
Girvan, Michelle and Mark EJ Newman (2002). “Community structure in social and biological networks”. In: *Proceedings of the national academy of sciences* 99.12, pp. 7821–7826.
Gleiser, Pablo M. and Leon Danon (2003). “Community Structure in Jazz”. In: *Advances in Complex Systems* 6.4, pp. 565–573.
Gomes, Dylan GE et al. (2024). “Marine heatwaves disrupt ecosystem structure and function via altered food webs and energy flux”. In: *Nature Communications* 15.1, p. 1988.
Gomez, Sergio et al. (2013). “Diffusion dynamics on multiplex networks”. In: *Physical review letters* 110.2, p. 028701.
Gouveia, Catarina, Ágnes Móréh, and Ferenc Jordán (2021). “Combining centrality indices: maximizing the predictability of keystone species in food webs”. In: *Ecological Indicators* 126, p. 107617.
Grime, J Philip (2006). *Plant strategies, vegetation processes, and ecosystem properties*. John Wiley & Sons.
Hanzl-Weiss, Doris, Mario Holzner, and Roman Römisch (2018). “Bratislava and Vienna: Twin Cities with big Development Potentials”. In: URL: `https://api.semanticscholar.org/CorpusID:169227102`.
Harmon, Jason P, Nancy A Moran, and Anthony R Ives (2009). “Species response to environmental change: impacts of food web interactions and evolution”. In: *Science* 323.5919, pp. 1347–1350.
Hidalgo, César A (2021). “Economic complexity theory and applications”. In: *Nature Reviews Physics* 3.2, pp. 92–113.
Hidalgo, César A et al. (2007). “The product space conditions the development of nations”. In: *Science* 317.5837, pp. 482–487.
Hidalgo, César A. and Ricardo Hausmann (2009). “The building blocks of economic complexity”. In: *Proceedings of the National Academy of Sciences* 106.26, pp. 10570–10575.
Holme, Petter et al. (2002). “Attack vulnerability of complex networks”. In: *Physical review E* 65.5, p. 056109.
IRiE (2022). `https://gis-portal.espon.eu/arcgis/apps/sites/#/irie-hub?`. Accessed: 2024-10-29.

IRiE Final Report (2022). https://gis-portal.espon.eu/arcgis/sharing/rest/content/items/5d1239c0336041fdb87ef00454f61bb9/data. Accessed: 2024-12-05.

Ishutkina, Mariya and R John Hansman (2008). "Analysis of Interaction between Air Transportation and Economic Activity." In: *The 26th Congress of ICAS and 8th AIAA ATIO*, p. 8888.

Jackson, Sharon and Stefan Markowski (1995). "The Attractiveness of Countries to Foreign Direct Investment". In: *J. World Trade* 29, p. 159.

Jazz musicians network dataset – KONECT (2017). Last accessed 06.2024. URL: http://konect.cc/networks/arenas-jazz.

Jordán, Ferenc, Andras Takacs-Santa, and Istvan Molnar (1999). "A reliability theoretical quest for keystones". In: *Oikos*, pp. 453–462.

Kang, Yuhao et al. (2020). "Multiscale dynamic human mobility flow dataset in the US during the COVID-19 epidemic". In: *Scientific data* 7.1, p. 390.

Karimi, Fariba and Matthias Raddant (2016). "Cascades in real interbank markets". In: *Computational Economics* 47.1, pp. 49–66.

Keyes, Aislyn A, Allison K Barner, and Laura E Dee (2024). "Synthesising the Relationships Between Food Web Structure and Robustness". In: *Ecology Letters* 27.10, e14533.

Keyes, Aislyn A, John P McLaughlin, et al. (2021). "An ecological network approach to predict ecosystem service vulnerability to species losses". In: *Nature communications* 12.1, p. 1586.

Kivelä, Mikko et al. (2014). "Multilayer networks". In: *Journal of complex networks* 2.3, pp. 203–271.

Kleinberg, Jon M. (1999). "Authoritative Sources in a Hyperlinked Environment". In: *J. ACM* 46.5, pp. 604–632. ISSN: 0004-5411.

Komornicki, Tomasz, Piotr Rosik, and Marcin Mazur (2023). *A multilayer vision of regional integration*. Retrieved on 09 Dec 2024. Luxembourg. URL: https://coilink.org/20.500.12592/m7czwk.

Krause, Ann E et al. (2003). "Compartments revealed in food-web structure". In: *Nature* 426.6964, pp. 282–285.

Kunegis, Jérôme (2013). "KONECT – The Koblenz Network Collection". In: *Proc. Int. Conf. on World Wide Web Companion*, pp. 1343–1350. URL: http://dl.acm.org/citation.cfm?id=2488173.

Lee, Kwang-Hoon (2016). "The conceptualization of country attractiveness: a review of research". In: *International Review of Administrative Sciences* 82.4, pp. 807–826.

Luczkovich, Joseph J et al. (2003). "Defining and measuring trophic role similarity in food webs using regular equivalence". In: *Journal of Theoretical Biology* 220.3, pp. 303–321.

Lusseau, D. et al. (2003). "The Bottlenose Dolphin Community of Doubtful Sound Features a Large Proportion of Long-Lasting Associations". In: *Behav. Ecol. and Sociobiol.* 54, pp. 396–405.

Mariani, Manuel Sebastian et al. (2019). "Nestedness in complex networks: observation, emergence, and implications". In: *Physics Reports* 813, pp. 1–90.

Martínez-Ramos, Miguel et al. (2016). "Anthropogenic disturbances jeopardize biodiversity conservation within tropical rainforest reserves". In: *Proceedings of the National Academy of Sciences* 113.19, pp. 5323–5328.

McDonald-Madden, Eve et al. (2016). "Using food-web theory to conserve ecosystems". In: *Nature communications* 7.1, p. 10245.

Mealy, Penny, J. Doyne Farmer, and Alexander Teytelboym (2019). "Interpreting economic complexity". In: *Science Advances* 5.1, eaau1705.

Miller, Ronald E and Umed Temurshoev (2017). "Output upstreamness and input downstreamness of industries/countries in world production". In: *International regional science review* 40.5, pp. 443–475.

Morris, Aaron L et al. (2016). "Deforestation-driven food-web collapse linked to emerging tropical infectious disease, Mycobacterium ulcerans". In: *Science Advances* 2.12, e1600387.

Musolino, Dario (2016). "L'attrattività percepita di regioni e province del Mezzogiorno per gli investimenti produttivi". In: *Rivista economica del Mezzogiorno* 30.1, pp. 45–70.

Musolino, Dario and Balazs Kotosz (2024). "A new territorial attractiveness index at the international scale: design, application and patterns in Italy". In: *The Annals of Regional Science* 72.4, pp. 1159–1187.

Nagelkerken, Ivan et al. (2020). "Trophic pyramids reorganize when food web architecture fails to adjust to ocean change". In: *Science* 369.6505, pp. 829–832.

Newman, Mark (2018). *Networks*. Oxford university press.

OECD (2022). *Measuring the attractiveness of regions*. `https://doi.org/10.1787/fbe44086-en`. Accessed: 2024-12-05.

Oldham, Stuart et al. (2019). "Consistency and differences between centrality measures across distinct classes of networks". In: *PloS one* 14.7, e0220061.

Öner, Özge (2017). "Retail city: the relationship between place attractiveness and accessibility to shops". In: *Spatial Economic Analysis* 12.1, pp. 72–91.

Opsahl, Tore (2011). *Why Anchorage is not (that) important: Binary ties and Sample selection*. http://wp.me/poFcY-Vw , Last accessed 12.2024. URL: http://wp.me/poFcY-Vw.

Otto, Sonja B, Björn C Rall, and Ulrich Brose (2007). "Allometric degree distributions facilitate food-web stability". In: *Nature* 450.7173, pp. 1226–1229.

Page, Lawrence (1999). *The PageRank citation ranking: Bringing order to the web*. Tech. rep. Technical Report.

Parisi, Federica, Tiziano Squartini, and Diego Garlaschelli (2020). "A faster horse on a safer trail: generalized inference for the efficient reconstruction of weighted networks". In: *New Journal of Physics* 22.5, p. 053053.

Peri, Giovanni and Francisco Requena-Silvente (2010). "The trade creation effect of immigrants: evidence from the remarkable case of Spain". In: *Canadian Journal of Economics/Revue canadienne d'économique* 43.4, pp. 1433–1459.

Petchey, Owen L et al. (1999). "Environmental warming alters food-web structure and ecosystem function". In: *Nature* 402.6757, pp. 69–72.

Pillai, Pradeep, Andrew Gonzalez, and Michel Loreau (2011). "Metacommunity theory explains the emergence of food web complexity". In: *Proceedings of the National Academy of Sciences* 108.48, pp. 19293–19298.

Polis, Gary A (1991). "Complex trophic interactions in deserts: an empirical critique of food-web theory". In: *The American Naturalist* 138.1, pp. 123–155.

Provenzano, Davide, Bartosz Hawelka, and Rodolfo Baggio (2018). "The mobility network of European tourists: a longitudinal study and a comparison with geo-located Twitter data". In: *Tourism Review* 73.1, pp. 28–43.

Pugliese, Emanuele, Andrea Zaccaria, and Luciano Pietronero (2016). "On the convergence of the Fitness-Complexity Algorithm". In: *The European Physical Journal Special Topics* 225.10, pp. 1893–1911.

Rauch, James E and Vitor Trindade (2002). "Ethnic Chinese networks in international trade". In: *Review of Economics and Statistics* 84.1, pp. 116–130.

Reiner, Christian, Susanne Meyer, and Sascha Sardadvar (2017). "Urban attraction policies for international academic talent: Munich and Vienna in comparison". In: *Cities* 61, pp. 27–35.

Rocchi, Marta et al. (2017). "Key species and impact of fishery through food web analysis: a case study from Baja California Sur, Mexico". In: *Journal of Marine Systems* 165, pp. 92–102.

Rodgers, Niall, Peter Tiňo, and Samuel Johnson (2023). "Influence and influenceability: global directionality in directed complex networks". In: *Royal Society Open Science* 10.8, p. 221380.

Rosvall, Martin, Daniel Axelsson, and Carl T Bergstrom (2009). "The map equation". In: *The European Physical Journal Special Topics* 178.1, pp. 13–23.

Russo, A et al. (2013). "ATTREG-The attractiveness of European regions and cities for residents and visitors". In: *Applied Research* 1.7.

Santana-Gallego, María, Francisco Ledesma-Rodríguez, and Jorge V Pérez-Rodríguez (2011). "Tourism and trade in OECD countries. A dynamic heterogeneous panel data analysis". In: *Empirical Economics* 41.2, pp. 533–554.

Schweitzer, Frank et al. (2009). "Economic networks: The new challenges". In: *science* 325.5939, pp. 422–425.

Sciarra, Carla et al. (2020). "Reconciling contrasting views on economic complexity". In: *Nature communications* 11.1, p. 3352.

Scott, Allen J (2001). *Global city-regions: trends, theory, policy*. OUP Oxford.

Servedio, V. D. P. et al. (2005). "Community structure from spectral properties in complex networks". In: *AIP Conference Proceedings* 776.1, pp. 277–286. ISSN: 0094-243X.

Servedio, Vito D. P. (2024). *A non-optimized Julia code to estimate fitness centrality in networks*. Version 0.1.0. `https://github.com/vitelot/FitnessCentrality`.

Servedio, Vito Domenico Pietro et al. (2018). "A new and stable estimation method of country economic fitness and product complexity". In: *Entropy* 20.10, p. 783.

Servillo, Loris, Rob Atkinson, and Antonio Paolo Russo (2012). "Territorial attractiveness in EU urban and spatial policy: A critical review and future research agenda". In: *European urban and regional studies* 19.4, pp. 349–365.

Squartini, Tiziano, Rossana Mastrandrea, and Diego Garlaschelli (2015). "Unbiased sampling of network ensembles". In: *New Journal of Physics* 17.2, p. 023052.

Stangl, Johannes et al. (2024). "Firm-level supply chains to minimize unemployment and economic losses in rapid decarbonization scenarios". In: *Nature Sustainability*, pp. 1–9.

Straccamore, Matteo et al. (2023). "Urban economic fitness and complexity from patent data". In: *Scientific Reports* 13.1, p. 3655.
Tacchella, Andrea, Matthieu Cristelli, et al. (2012). "A new metrics for countries' fitness and products' complexity". In: *Scientific reports* 2.1, pp. 1–7.
— (2013). "Economic complexity: conceptual grounding of a new metrics for global competitiveness". In: *Journal of Economic Dynamics and Control* 37.8, pp. 1683–1691.
Tacchella, Andrea, Dario Mazzilli, and Luciano Pietronero (2018). "A dynamical systems approach to gross domestic product forecasting". In: *Nature Physics* 14.8, pp. 861–865.
Taylor, Peter and Ben Derudder (2015). *World city network: a global urban analysis*. Routledge.
Thurner, Stefan, Rudolf Hanel, and Peter Klimek (2018). *Introduction to the theory of complex systems*. Oxford University Press.
Ulanowicz, Robert E, Cristina Bondavalli, and Michael S Egnotovich (1998). "Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 97: The Florida Bay Ecosystem". In: *Annual Report to the United States Geological Service Biological Resources Division. Ref. No. [UMCES]CBL*, pp. 98–123.
Ulanowicz, Robert E and Charles J Puccia (1990). "Mixed trophic impacts in ecosystems". In: *Coenoses*, pp. 7–16.
US airports network dataset – KONECT (2017). Last accessed 12.2024. URL: http://konect.cc/networks/opsahl-usairport.
Vallarano, Nicolò et al. (2021). "Fast and scalable likelihood maximization for exponential random graph models with local constraints". In: *Scientific Reports* 11.1, p. 15227.
Waltert, Fabian and Felix Schläpfer (2010). "Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies". In: *Ecological Economics* 70.2, pp. 141–152.
Weng, Jianshu et al. (2010). "Twitterrank: finding topic-sensitive influential twitterers". In: *Proceedings of the third ACM international conference on Web search and data mining*, pp. 261–270.
White, John G. et al. (1986). "The Structure of the Nervous System of the Nematode Caenorhabditis elegans". In: *Phil. Trans. R. Soc. Lond* 314, pp. 1–340.
World Bank and KNOMAD (2023). "Leveraging Diaspora Finances for Private Capital Mobilization". In: *Migration and Development Brief* 39. URL: https://https://documents1.worldbank.org/curat

ed/en/099740408142422676/pdf/IDU-84dfd61b-e135-4242-a202-3728b2e8fa86.pdf.
Yodzis, Peter (1998). "Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem". In: *Journal of Animal Ecology* 67.4, pp. 635–658.
Zachary, Wayne (1977). "An Information Flow Model for Conflict and Fission in Small Groups". In: *J. of Anthropol. Res.* 33, pp. 452–473.
Zachary karate club network dataset – KONECT (2017). Last accessed 06.2024. URL: http://konect.cc/networks/ucidata-zachary.